

## Figure 2 Methods for the diversification of DNA sequences coding subunits of heteropolymeric proteins and testing for superior variants.

The example given for immunoglobulins is for illustrative purposes only and is not intended to limit application of the present invention to this specific heteromeric protein. H = heavy chain genes, L = light chain genes

Figure 2A existing protocol: Number of transfections needed to generate 1024 new combinations: 2048

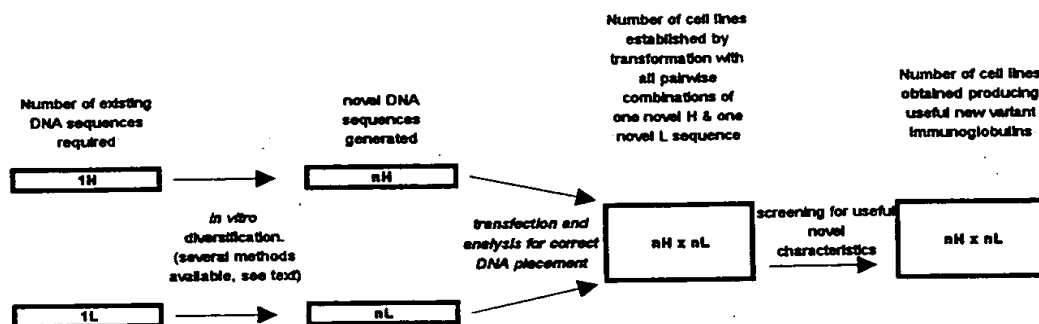


Figure 2B existing protocol using the heterokaryon technology of US Patent Serial No. 5,643,745  
Number of transfections needed to generate 1024 new combinations: 64

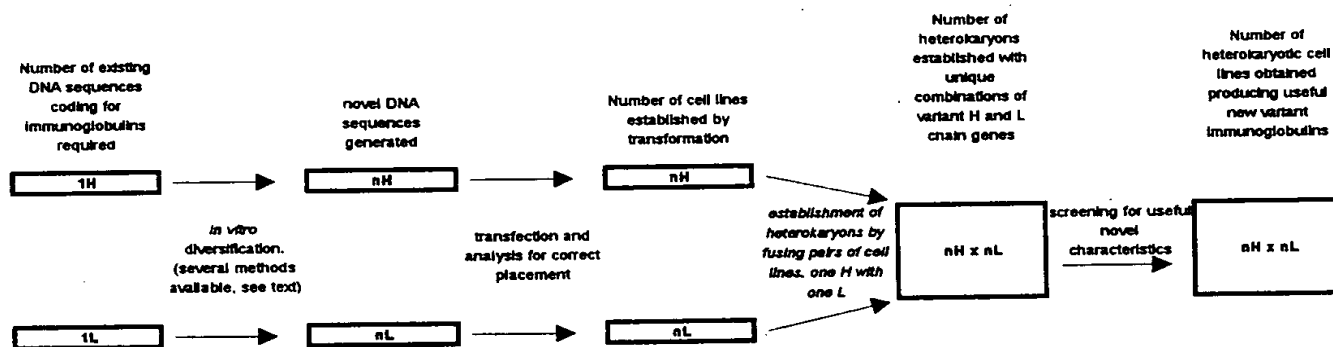
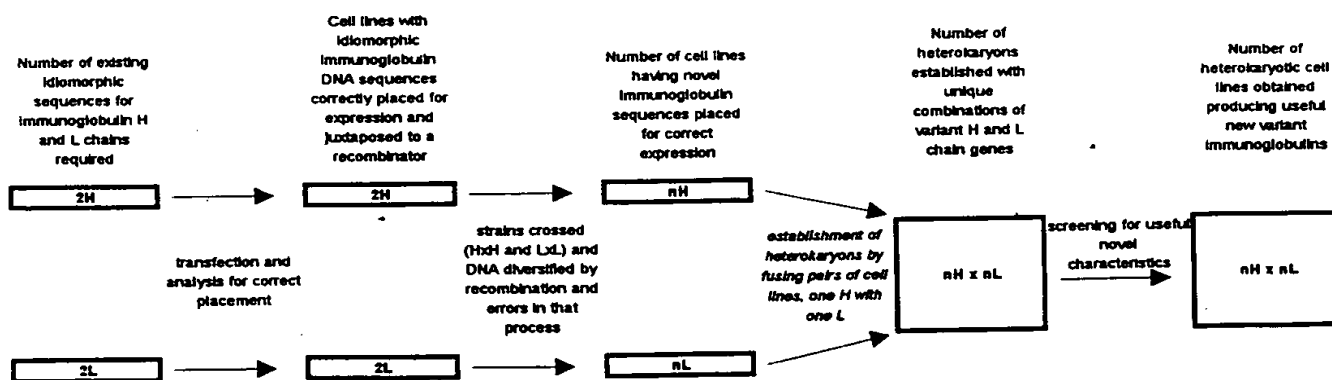
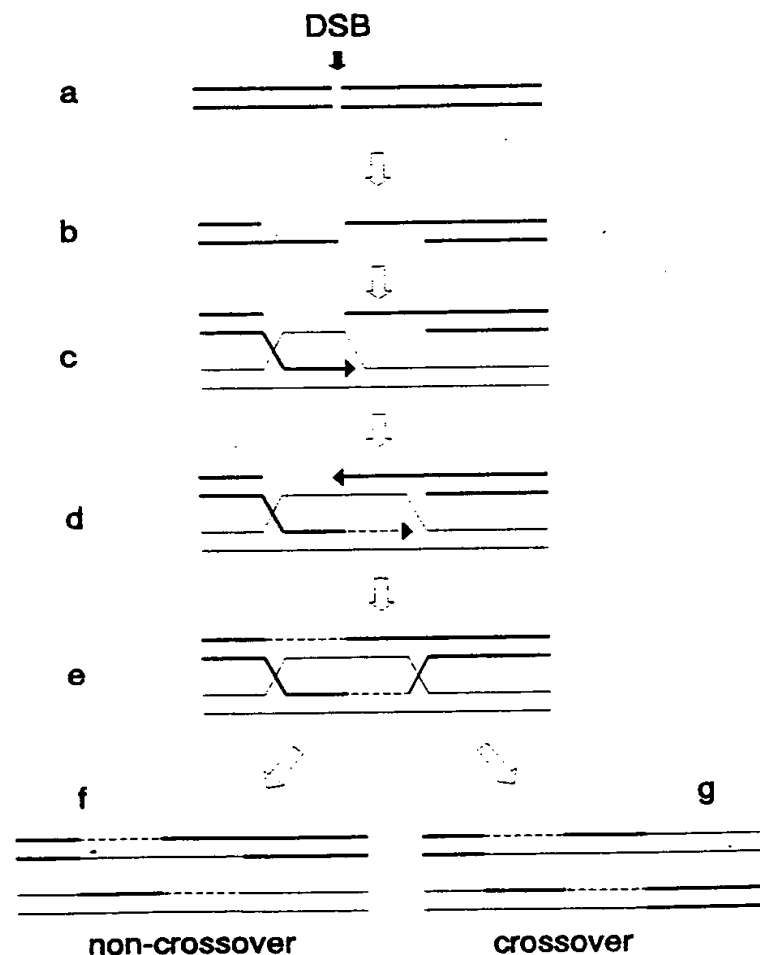


Figure 2C a protocol enabled by the present invention  
Number of transfections needed to generate 1024 new combinations: 4



**Figure 3** The modified double strand break repair model for meiotic recombination. After H Sun *et al* Cell **64**: 1155-1161, 1991

(a) A double strand break (DSB) is made in one DNA duplex. (b) A long 3' overhanging single strand tail is generated either side of the break by resection. (c) One 3' end invades a homologous duplex forming a D loop. (d) the D loop is enlarged by repair synthesis and anneals to the second 3' end. (e) Repair synthesis occurs at the second 3' end and two intermolecular junctions (Holliday junctions) are formed. Resolution of the junctions by cutting inner and outer strands can give rise to non-crossover (f) and crossover (g) chromosomes. If there are base mismatches in the heteroduplex regions (duplex molecules with thick and thin lines) there will be gene conversion. If mismatch repair does not occur there will be post meiotic segregation of new sequence combinations.



**Figure 4** Life cycle of *Neurospora crassa* after JRS Fincham (Genetics, Wright 1983). Microconidia having one nucleus are not shown but can be generated as described in the text. Perithecia and protoperithecia are shown in section.

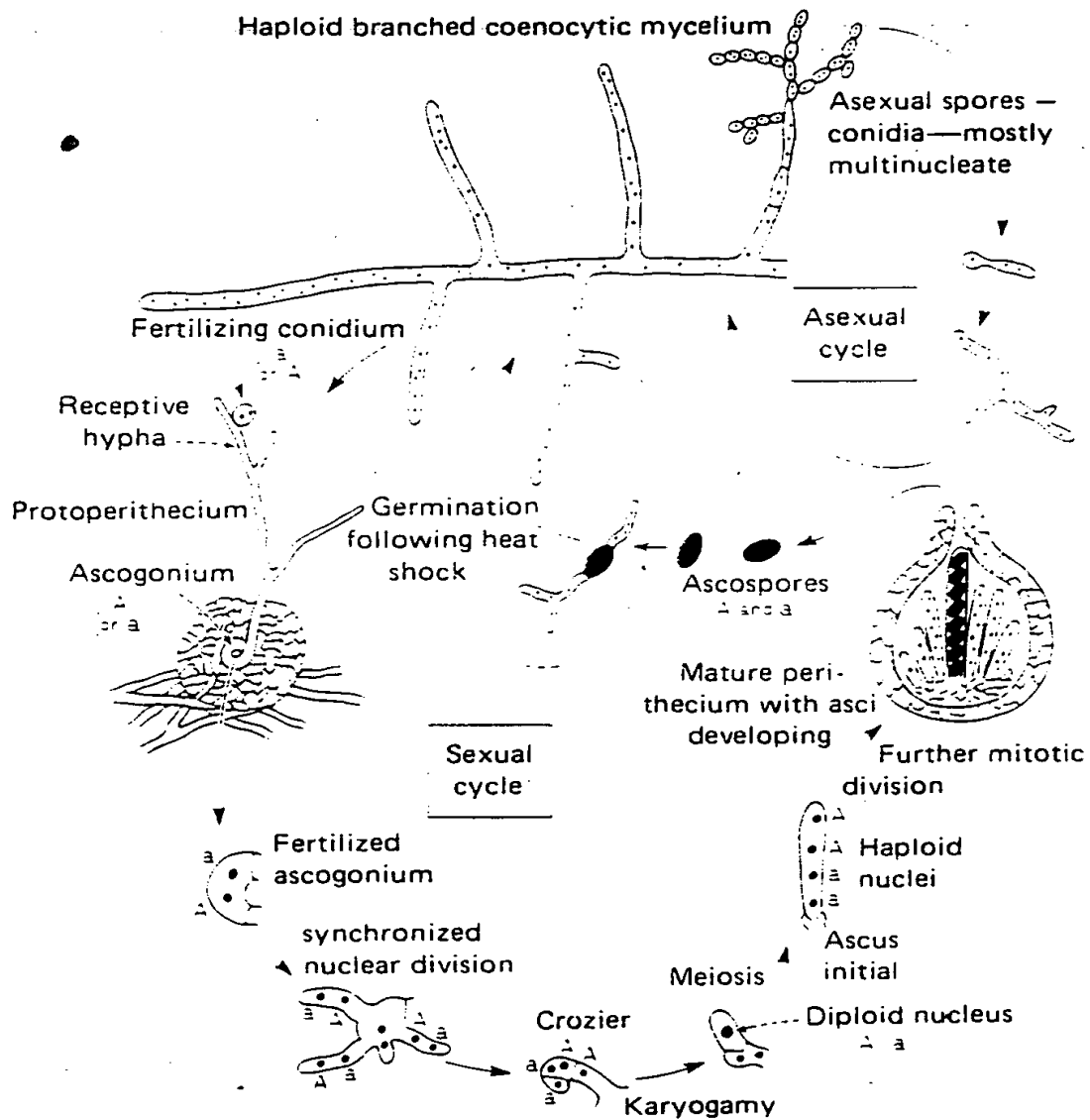




Figure 6. Discontinuity in the parental origin of DNA sequences in progeny from crosses between pairs of *his-3* alleles. In most cases this reflects discontinuity of conversion tracts, in some cases crossovers near the ends of conversion tracts.

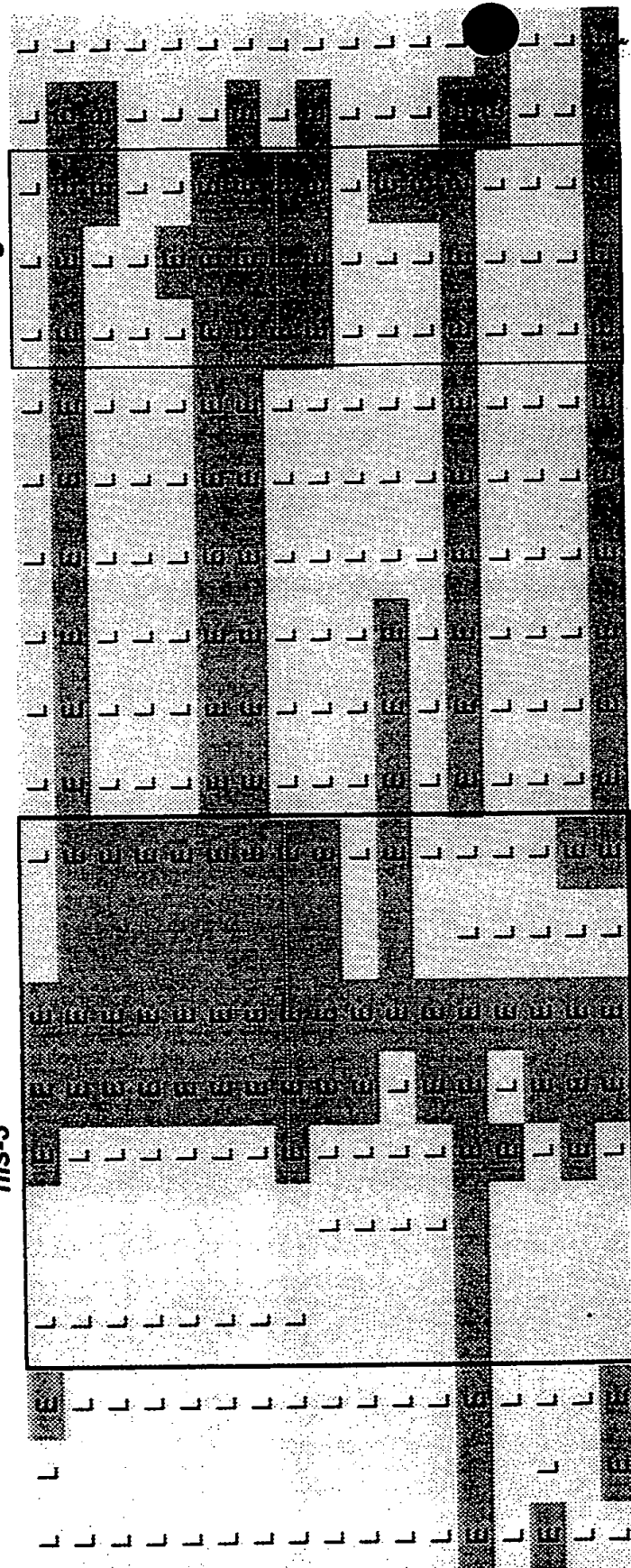
Markers are specific DNA sequence differences that distinguish the parents. These were all E (Emerson wild type origin) or all L (Lindgren wild type origin) in the parental strains. Recombinants carry both E and L markers.

Marker position is given in base pairs from the first base of the first codon of the *his-3* gene. Each line of the table shows the parental origin of the markers inherited by one of the progeny.

marker	P	H	P1	K1201	K504	L3	R1	K26	K874	R4	C4	C5	C6	C6/7	C7	C8	C1	C2	C3	C9	D
location	~600	-384	115	179	563	1232	1502	1717	2318	3436	3705	4000	4304	4667	4821	5232	5495	6153	6507		

*his-3*

*cog*



**Figure 7** Nucleotide sequence of the *his-3 cog<sup>L</sup> lpl* region of linkage group I in the Lindegren wild type strain of *Neurospora crassa*. This differs from that in the StLawrence strain in many positions, summarised in figure 5. The coordinates of relevant features are given in the text. This sequence contains the high frequency recombinator *cog<sup>L</sup>* which is active providing the cross in which meiosis occurs is homozygous *rec-2*.

```

1  GATCGCAACT  GGAGATCACT  CGCACCGTGC  CGCAGAACAA  GGGCGACGAG  CCTCAGGGCA
61  GTTTAGCCTG  CCGTAACAGC  ACAGACCATA  GCTTATTTTC  ACCTGGGCGG  GCGGGCGACG
121  GCGGCACTGA  CATCGGCAAG  GCGGCATCAA  GCAACCCCTC  TGTTGCTTGC  CAGCTGCCGG
181  CCAACGTCAG  CCGTACAAGG  AGAAATCTGG  AAGGAAAGAC  TTCTGGCACC  GACAGGATGG
241  CACGCGGGAA  AAGTTCCCAA  TGCATGAGAT  GAGGGGCATT  TGCATTGCCT  CCCGTCACAC
301  TGCCCCGCAA  CCCCACCCG  ACCATAGCGT  CTGTCGATAC  ATGGAGCGCG  AAGTTCAGAA
361  ACCTGTAATT  CCTGGTAACT  TTCAGGTACA  CAGTACGTAC  TGATCCTGGT  ATCAAACCTT
421  GCCTGCCGAG  TTTTCGACGG  AAAGAGGTGT  GAATTGTGAA  AGAGTCATAC  CAAATCACCC
481  GATTTTCATA  AAGCCCGAGT  CTTTCTGTGA  CATAAGCGAC  ACTCGAAGCG  GGCCTCATCT
541  TCATAGCCTG  ATAGCTTGTA  ATACTCCATC  CTCGTATCTC  ACTTGACCTT  GAGTTCAACC
601  CCACGTCAGA  CTTACCCGA  CACATCGACG  GATTGGGGAA  CAGCACAAAT  CCTGAAAAGC
661  GAGAAAACCA  AACAGAGGAA  AACACCATGG  AGACAACACT  TCCCCTCCCC  TTCCTCGTCG
721  GTGTCAGTGT  TCCTCCCGGA  CTGAATGACA  TCAAGGAGGG  CCTCAGCCGG  GAGGAAGTCT
781  CGTGTCTTGG  CTGCGTCTTC  TTCGAGGTCA  AGCCCAAGAC  CCTTGAGAAA  ATCGTGCGAT
841  TCCTCAAGCG  TCACAATGTC  GAATTGTAGC  CCTACTTCGA  TGTAAACAGC  CTCGAGTCTA
901  TCGATGATAT  TATCACTCTT  CTGGACGCGG  GCGCCCGCAA  GGTGTTTGTC  AAGACCGAGC
961  AGTTGGCCGA  CCTCTCCGCA  TATGGCTCCC  GCGTTGCCCC  CATTGTCACT  GGAAGCAGCG
1021  CTGCTTTGCT  TTCCTCCGCC  ACCGAGAGCG  GCCTTTTGCT  CTCCGGCTTC  GATCAGACTG
1081  CCTCCGAGGC  TGCACAGTTT  CTGGAGGAGG  CCAGAGACAA  GAAATATTAC  CCCTTCTTCA
1141  TCAAGCCCGT  TCCTGGGGCC  GATCTCGAAC  AGTTCATCCA  GGTGTTTGTC  AAGACCGAGC
1201  CCATCCCCAT  ACTGCCATCC  CTGGCTTGA  CAACAAAGAA  GGACGAGGCC  GGAAGCTTGT
1261  CCATCTCCAC  CATCCTCTCG  AGCGTCTGGA  AGTCTGACCG  TCCCGATGGT  CTGCTCCCCA
1321  CCGTTGTCGT  TGATGAGCAC  GACACTGCTC  TGGGTCTGGT  CTACAGCAGT  GCCGAGAGTG
1381  TGAACGAGGC  CCTCAGGACA  CAGACTGGTG  TCTATCAGAG  CCGGAAGCGC  GGTCTCTGGT
1441  ACAAGGGTGC  TACTTCCGGA  GACACTCAGG  AGTCTGTCGG  CATCTCGCTT  GATCAGTCCG
1501  ACGATGCTCT  CAAGTTTGTC  TGAAGCAGA  AGGGTCGTTT  CTGCCACCTC  TCGAGGAAAC
1561  GCTGCTTTGG  TCAGCTCAAA  GGCCTTCCCA  AGCTCGAGCA  GACTTTGATT  CTAGTCCGGG
1621  AGTCTGCCCC  CGAGGGCTCC  TACACTGCCC  GTCTCTTCTC  CGATGAGAAG  GAAATCGCCT
1681  CCAAGATCAT  GGAGGAGGCT  GAGGAGCTCT  GCACCGCTCA  GACCCCCAG  GCGGTTACTC
1741  TTGAGGCTGC  CGATCTCTTC  TACTTTGCTC  TTACCAGGGC  CGTTGCTGCC  AGGACTGGAG
1801  TTGCCGATAT  CGAAAGGAGC  CTTGACGCCA  AGAGCTGGAA  GGTCAAGCGC  GCTCCCGCTG
1861  ATGCTAAGGG  TAAGTGGGCT  GAGAAGGAGG  GCATCAAGCC  TGCGGCGTCC  AAGATCACCA
1921  CCACTTCGGC  CCCTGTCACC  AAGGAGGCCG  CCCAGGAGAC  CACCCCTGAG  CGATGCTGCT
1981  TGAGACGTTT  CGACGCCTCC  AAGGTCTCTA  CCGAGGAGCT  CCCCATCATC  GAGGACGTCC
2041  CTGCGCAAAA  GTCGTCCGAT  GTCATCTACA  AGATCATTGT  GTTCGAGAAG  GCTACCTCTC
2101  GCAAGAACGG  CGACAAGGCT  GTTCTGTCTG  ACACTCACAA  GTTCGAGACT  CCTGAGGAGA
2161  TTAGTAGCCC  CGTCCTGAAG  GCGCCCTTCC  CCAAGGAGCT  TATGCAGCTC  GCCCAGAAGG
2221  CCATTGCTGC  CATCGACGTG  TCCTTCGAGA  ACATCCGCAA  GTTCCACGCC  TTCTCTCGTC
2281  AGGAGAAGCC  CCTCCAGGTC  GAGACCATGC  CCGGTGTTGT  CTGCAGCCGT  AGCACTGCCC
2341  CCATCGAGGC  CGTCGGCTGC  TACATCCCCG  GCGGTACCGC  CGTTCTCCCC  GCCTCTCCTC
2401  TTATGCTGGG  TGTTCGCGCC  ATGGTCGCCG  GCTGCAACAA  GATTGTGTTT  GTTGGGGCCG
2461  CCCGCGCCGA  CGGAACCATC  ACTCCCGAGA  TTGTCCACGT  CGTCAACAAG  CATGGCCTAC
2521  AGTCCATCGT  GCTTGCCGGC  GGTGCCAGG  CCGTAGCTGC  CATGGCCTAC  GGTGCTCACT
2581  GCATCACCAA  GGTCGACAAG  ATTCTCGGCC  CCGGTAACCA  GTTCGTCACT  GGTGCTCACT
2641  TGTTCTGTCAG  CAACGACACC  AACGCTGCCG  TTGGGATTGA  CATGCCCCGT  GGTGCTCACT
2701  AGGTGCTGGT  CATCGCTGAC  AAGGACGCCA  ACCCCGCGTT  CGTTGCCTCG  GGTGCTCACT
2761  CCCAGGCTGA  GCACGGCGTT  GACAGTCAGG  TCATCTGAT  CGCTATTAAC  CTCGACGAGG
2821  AGCATCTTCA  GGCTATTGAG  GACGAGGTTT  ACCGTCAGG  TATGGAGCTT  CCTCGGCTCC
2881  AGATTGTCCG  TGGCTCCATC  GCCCACTCGA  TCACCGTGCA  GGTCAAGACC  GTGAGGAGG
2941  CCATGGAGCT  CAGCAACAAG  TACGCTCCTG  AGCACTTGAT  CCTCCAGATC  AAGGAGGCCG
3001  AGAAAGCTGT  CGATCTTGTC  ATGAACGCTG  GTAGTGTCTT  CATTGGCGCT  TGGACTCCTG
3061  AGTCCGTTGG  CGATTACTCT  GCTGGTGTTA  ACCACTCGCT  GCGTAAGTTA  CATATCATAA
3121  ATAGCCCCGC  TTCACAGATT  CTTCTGCTAA  CGTCAAGACA  CATAGCTACC  TATGGTTTTG
3181  GCAAGCAGTA  CTCTGGCGTC  AATCTCGCCT  GCTTCGTCAA  GCACATTACC  AGCTCCAAC
3241  TGAAGTCCGA  GGGTCTCAAA  AACGTCGGCC  AGGCTGTCAT  GCAGTTGGCT  AAGGTTGAGG
3301  AGCTCGAGGC  TCACAGAAGG  GCGGTCAGCA  TCCGTCTTGA  GCACATGAGC  AAGAGCAACT

```

Figure 7 continued

3361	AGACGGAAAT	TCTTTTTCGA	AGTTGCAAAA	AAAACAAGAA	CAAAGGATG	TAGTGGGTTG
3421	ATGTATATCT	GGGTCAATTT	GGGCACATAG	AGTAATGATA	ACGAGTTTTG	GACATTGTAC
3481	TGTTCTGTAC	AGGCTGAAGA	TCAGTACATG	AATCTGTTGG	TAAGTGTAGA	GACCCAAACG
3541	TCCCTTGAGT	TTTTCTCCCT	GTTCCAGAGA	GGTGCTCGTC	CCTGGGTGTT	TATTTTCATT
3601	ATTACATCAA	CCTTTTATTT	TATTTTATTT	TTTATTTTAC	TTTTTTTTTC	TTTTTTTCAG
3661	ATCATGCGTA	CATGAACGGG	GGAAGCACAG	ACGATCGAAA	CGTGGATGTC	ACAATGTCGC
3721	TGCAGTGATG	CTGCATTGCA	TGAAGCGCCC	ATCTCAATAT	ACTTGCAGTC	TTGCGCGTTG
3781	CACGTGAAC	TCCCAAACAA	CCGAATAAAA	GACGGCGAAA	AATGAAGATA	AAAAAAAACC
3841	ATAATAAAAA	TCGGAGGGAG	TGTGGGAAAT	GGTTTCTTTT	AGCATTTAGA	CCCCATAGCC
3901	GTGCACGCCC	GGGTACAGAC	AGGTTTCATCG	ATGTTGACAT	TGACTGGGAC	ACCAGGTCTA
3961	TCTATTTTCT	CTCCTGTCTT	CTACCATACA	TCGGGACATC	GGACATCTCG	CTGTACCCCC
4021	CACACCCACA	AAGTCTTATA	AAAGCGCCAC	ACCCGAGGAG	GTTGCGTCGG	CCCCACGAAC
4081	TCCGTGCCCT	CCTGCCTGTT	TACAGGGACC	GAACGCTGGA	GAAGCTTAGT	TTCCTGACAT
4141	CCGGCCTACC	CGAGCAGGAA	AAGGGACAGC	TCATAGGCGA	GGAGGGATTT	GAAGATGGGG
4201	ACATTTTGGA	TGATTTCGAGA	GGAGGAACTA	GGTACTGTAT	CATGATAGTT	CGGGGCAGCA
4261	TCTTGCTGG	GACATTGTTA	ATACCTCGAT	ATGATGAAGT	GGGAGGGAGT	TTTTTCATGT
4321	CTTGCCCAAG	TCCCACTAAT	CTTTTTTTTT	TTTTGTACCA	ACACCCAAGA	TTCGGAGAAT
4381	AGTGTAAAGGA	TTGCGATTCA	CAAGTGGAAG	TCTGAGGATC	TTTTTATATC	TTTGTCTTCC
4441	GCGGACTGTT	AACGATCCTA	CAGCGAGCGA	GCGAGCGGTC	GGATGCGCTG	ATCTGATAGG
4501	TGCAATATAC	GGCCGCTTTC	TCCGGTCGTC	TAGTGTAAGC	TCTGTCCGCA	TAGTAGTACA
4561	CTAAAAAACC	CCTTGCAATT	CATGATCTGC	TTGCTATTCA	TTCCGAGTTA	TTTCAGTGGT
4621	CACATTTTCA	GATTCACAGC	CATCCATCCA	TATGGAAAAA	TCCATTCCCA	TGCTTCCCTC
4681	CCCCCACTAT	GTATGTGACC	ACACGCTGCT	GTCAGAATGC	CAACGGTCTC	AGGTACCCTC
4741	GTCCGACTGT	TTGGCATGGA	GTTACATACA	CTACTAGTGT	AGCCCCGGGC	CAAGCTACCC
4801	CGTCAAATCT	ATACATATCT	ATAATGGGTT	TCAGGTGTTT	CGTTCGCTGT	CAATCAAGTT
4861	TGAAACATCA	CTGGGGCCGT	TGGACGGTGT	ATTAGACCAT	TGGCTCCCTC	AGCTGGCGGC
4921	TGGGCGGTTG	GGTCGGCAAT	AACGGGACTG	GACTTGAGAG	GGACGAGGAG	AGTCGGTTGG
4981	CTGCCTACAC	TACACTACAA	GCGTTCCAC	CTAACCGACG	AGTCCCGTTT	TCCATTTGTG
5041	TGCCTTAACC	ATCATCTAGG	GATGTCAGGG	TTTGGCCGGA	TCAGGGTATG	TTTGTTGAC
5101	TGTTGTCTAT	TCTGATTGGG	TACATATCAT	GGTAGGTGTC	TCGAGAACAG	TAGAGTACTC
5161	GGGCTTAGCG	TTTGATGAT	TACGCGAGAT	ATGAGTTGTA	GGCCGCCATG	CAGTTGCTTG
5221	CCCATAAGCA	GAAGTTGCTT	TGGGATATAT	TTCTCGTCTT	TCAAAGGTCA	CGAGGTCCTG
5281	GGACGAGCGG	CATCGCCATC	CAAAGGGTTG	AACATGAGAA	ACCGGAATGG	CCTTTGCGTT
5341	GAAATACAAA	AAGTCAAGAA	TAAATCGCT	TGAGGATAGG	GACGTGGAAG	CAAGCAAATA
5401	TGGTAAGGGA	GGTACTGCTA	TGTAGGTGCT	CAGCAAACTG	CCAATTTCTT	GGCCCCCAAG
5461	CAGCAGTTTG	CTGTCAGTGC	TGTCGTGTC	AGCCTTGCTA	GTGGAACCTA	AACTGCTAAC
5521	ACAGCGCAAG	TGCGCATGTA	AAGATATTGT	GGGAGGATCT	GTATGGATGG	ATGAGATTAC
5581	TGCTTGGTGT	TGGTTGCGAG	GCACTGCGGC	TGTTAGGCTT	TGCTGTGCCC	CGTTTCGACG
5641	AGAAATACGC	GGAACATAAA	ATTGGATACC	TAGACTTACT	GCCTATGGGA	GGTATCTACC
5701	GACGTAGCCG	ACGGATTCTA	GCAACATCCC	GACTTTGCTT	GTAGTGTACT	ATGATAGCAG
5761	CACAGTGGGG	TGTTGCTCCT	TGTGAGCATG	GGCTCTTTTT	TTTTTTTTTC	CCCTTCCCTA
5821	GGGCGTTGAC	TSGACTTGCT	CTATCGTTCC	CAAGGTAGGT	GCCCCTCATC	GATTTTCCCA
5881	AGCCGTCTCC	CGCCAGATTG	TCGTTCATAG	GTCATGATGA	CCTCGGTCGC	TGGGGTGGC
5941	TGGTTACGGG	GAGCTGGGAC	CGCTAGGCC	CAGTGGTTGT	GCCATTGAGC	TGGGGTGTGT
6001	GGAGTACGGG	TAGAGGCGCT	TGGAAGTTGT	GCTAGCGGAA	ACCCTGGAAT	ATCTTGTACC
6061	CTTCGATTCC	TTCTCGGGCT	GCCCATGTGC	TGAGGTGATG	CCGGGGATCT	GGCGCCAATC
6121	ATCCATTGAG	GTTCCCGCAG	CTTCCCGGTG	CCGCGCGCGG	GCGCAGTTGC	TCACAGGACA
6181	CACCTAGACG	CAGGGGCACA	GGGGCACCGT	TTGGTGTGCA	ACTGGGTACC	TGGTAGCTGT
6241	AGCAAGCACT	CCACCGTCTG	TGCAATCCCC	CAATCCACGG	CAGGAACCTA	GCACCGCCGC
6301	GGCACCAGT	GAGCGAATCC	ATCCGCATTG	GATCCCAATT	CTTGCCCTTG	CCATCCTTCT
6361	TTCTTCCAC	TTGGCGCAAC	CAACACTTCC	CTTGGTCTGG	GTACTCGTGT	TGATCTTCAC
6421	TCTCTTTTTT	TCTTGGGCGA	CCGACTTTT	ATATCCGTCC	TTGCTTCCCC	CTGGCCGTTG
6481	TCGTTCTTTC	TACAACTACC	TTCCGTTTCT	TATCCCCTTT	CTTGGTTCCG	TCGAGGACCC
6541	AAAAACAGAA	CAATTCCGGC	TCTTCCAGGT	GGCTTGGGTG	CGACTGTTTA	GCTCTTGACC
6601	ACTAGCCGCT	TACCTTCTCT	TGATGTTTAT	ATTTGGATAT	CATTGAACCTA	CTCTTCTTGT
6661	AAACGGCAGA	CGAACGGAAC	AGTCCCTACG	GTTTATTAGC	GATATACGTT	GTACTGATAT
6721	CCTGAGCAAG	AAGAGGCAAA	TTATCAATTA	TGCATCTCCC	ATCGTCCGTC	CTCATCGCAG
6781	CTCCCTTGCT	CGCCAATGTA	TCGGCCGAAC	CGATTAGGAT	ACCCCAACGC	GATGTTCTCC
6841	GTGGTATCAA	CATCACAGCA	ACTTGCCGTT	CGAGCACTAC	CGAATTCGCC	CAGCGGTGGA
6901	TATGCCCTTG	CCGTTGTAGA	CTGTCCCAAG	ACCAAGCCGA	CGCTCCGGAA	GGCCGTGGAT
6961	TTGTGCAACG	AGGAGAAGAA	CTGGTTGTCT	ATCCGGAGGA	AGAACACCAT	CCAGCCCATG
7021	AGGGACCTAC	TGAAGAGGGC	CAACATCACT	GGGTTCGATT	CCGAAACTTT	CATGAATGAG
7081	GCCGCCAACA	ACGTCTCGCA	ACTGCCCAAT	CTCGCCATTG	GCTATTCAGG	ACGGGGCTAT
7141	CGTGCCCTCA	TGAACGGCGC	CGGCTTCGTT	GCTGCTGCGG	ATAACCGGAT	TCAAAATACC
7201	ACGGGCGCAG	GTGGTATTGG	AGGCTTGTTG	CAGTCCAGCA	CATATTGTGA	TGTAAAACCA
7261	TGCCTTCTTG	TGGTTCTTCT	TATCTCGTTT	TCGAGTGTCA	ACTGCGCCAG	TTCGACGTTG

Downloaded from www.sciencedirect.com

Figure 7 continued

7321	GGCGGCTGTG	GACGACCTTG	CTGGTGAACA	TGTCTTGGAC	TCCATGCCCC	TTTTTTCGTT
7381	CCCTAAAATC	CCAAAAAATA	AAAAAATAAA	AAAAAATAAA	AAAAAATAAA	AAAATTTCGAG
7441	GACCGTGACT	GTAATTTGCT	AACGCAACTC	TAGGGCCGGA	CTTTCTGGTG	GTGGCTGGCT
7501	TGTCGGCAGT	TTGTTCTCCA	ACAACCTTCAG	TAGCATTGAG	ACCCTGTGTA	GCGAGAACAA
7561	AGTCTGGGAC	TTTGAGAACT	CCATCTTTAA	AGGACCCAAG	GAGGCTGGCC	TTAGTACTGT
7621	CAACCGTATC	CAGTACTGGT	CCGAAGTGGC	AAAGGAAGTT	GCGAAGAAGA	AGGATGCTGG
7681	CTTCGAGACA	AGTATAACAG	ACTACTGGGG	CCGAGCATTG	AGTTACCAAC	TGATCGGAGC
7741	CGATATGGGC	GGCCCGGCTT	ACACCTTCTC	CAGCATTGCC	CAGACCGACA	ACTTCCAGAA
7801	GGCCGAAACG	CCGTTCCCTA	TTCTGGTAGC	TGACGGCCGC	GCGCCTGGAG	ACACCATCAT
7861	CTCCCTCAAT	GCTACCAACT	ACGAGTTCAA	CCCGTTCGAG	ACGGGTAGCT	GGGACCCGAC
7921	CGTCTATGGC	TTTGCGCCGA	CCAAGTACCT	CGGCGCCAAC	TTCAGCAACG	GCGTGATCCC
7981	ATCGGGAGGC	AAGTGC GTTG	AGGGTCTCGA	CCAAGCCGGC	TTCGTCATGG	GCACCAGCAG
8041	CACGCTCTTC	AACCAGTTCC	TTTTGGCCAA	CATCTCCAGC	TACGACGGTG	TTGCCAGACG
8101	TGCTCATCGA	GGCCGTGACT	TCTGTCTCTA	AGGAAATCGG	CGCCAAGAGG	ACGACGTCTC
8161	CCAAATCATC	CCTAATCCGT	TCCTGGACTG	GAACAACCGG	ACCAACCCCA	ACGCCGACAC
8221	GCTCGAGCTC	GACCTGGTCG	ACGGCGGCCG	AGATCTGCAG	AATATTCCGC	TCAACCCGCT
8281	CACCCAACCC	GTGCGCGCCG	TCGACGTCAT	CTTCGCTGTC	GACTCGTCCG	CCGACGTGAC
8341	AAACTGGCCC	AATGGCACCG	CCCTGCGCGC	CACCTACGAG	CGCACTTTTCG	GCTCTATTTT
8401	CAACGGGACA	CTCTTCCCTT	CGATCCCCGA	CGACTGGACG	TTTATAAACC	TAGGCCTCAA
8461	CAACCGCCCC	TCTTTCTTCG	GCTGCGATGT	TAAGAACTTT	ACCTTGAACG	CCAACCAAAA
8521	GGTTCCCCCC	TTAATCGTCT	ATGTCCCCAA	CGCGCCCTAT	ACCGCGCTGA	GCAACGTGTC
8581	CACCTTCGAT	CCGTCATACA	CGATGTCTCA	GCGCAACGAC	ATCATCGGCA	ACGGATGGAA
8641	CTCAGCCACG	CAGGGAAACG	GCACGCTGGA	TTCGGAGTGG	CCCACTTTCG	TCGCCTGCGC
8701	GGTTATCAGC	AGGAGCTTAG	ATCGGTTGGG	CAGGCAGACG	CCAGCCGCGT	GCAAGACTTG
8761	CTTTGACAGG	TATTGCTGGA	ATGGCACAGT	GAACTCCAAA	GATACGGGGG	TTTACATGCC
8821	TGAGTTCAAG	ATTGCGGATG	CGCATGCCCT	GGACTCGGGT	GCTGTTGCTA	TCGGAAAGAT
8881	GGTGAATGTC	TGGTCGTCGG	TTGTGGTGGG	AGTTGTGGCG	GCTACTTTGT	TGTTGTAGGG
8941	GTAGGGGAGA	CGTGATGATA	TTCCAGTCTG	ATGAAGTTGA	GACTGGACTG	GAGATCGCCA
9001	AGGATGCGGA	GGGAAAGGAA	TGCGTGGTGT	TAATGTCATG	ATGGATGAAG	AGTCATGGAT
9061	CATGGAACGA	CGGGGCGGGG	ATATTGGATG	ATGGATATAC	CACACTGCAT	GCATGCTCTA
9121	TTGATAGTAT	GCTTTGGCAT	TTACGTTTAA	CAATCAATTG	CTCCATCCTG	ATGTTCTATC
9181	TTTTTTCGACA	ATGGATTGAT	ACTACTCCTG	TTGCTTCGCT	CTTGAGGTTG	GAAGGACTTG
9241	AGGTTGGAAG	GACTTGAGGT	TGTTTGTTCT	GAGGGAGGTT	ATCGAAGTAT	CATCTGTGCT
9301	GATGCCGATT	GATAGACTGT	CCTCTTCTTC	GAGGCAACGA	ACGGTCGGAT	GAGCCTCTTT
9361	AATCATGATG	CTCAGTGCCA	CAAAAAGGCT	CCAGCACAGC	TGCCCACACC	TTTCTTGCCCT
9421	CGCCGTTTCT	TCCTTTTTCT	TTTCCCTGTG	TTCTTTTCTT	CCTTTCCATC	TCATCCCGTA
9481	CCAGAGTGCC	CACCGGGTAT	ATATATTACC	TCCTTGGCCG	TTCTCTTTTG	ACCAATAAAT
9541	CGCTTGGTCG	AGTGGCGTAA	CGGTTTACCG	TCTACACTTA	TCACTCAAAC	CAAACCAAAC
9601	CATCGAAGAA	GTGACCTATC	GGTTCGAGGG	AACGGTGATG	TTCTTACGAC	CAAGTTAACC
9661	CAAAGAGCGT	TCCACATCGT	TGAACCGTCT	CCTCCAGTTG	GATCTGTTTA	ACTTCCGCAG
9721	CGACTGAAGA	AGGTATCACT	TTTTTTTTTG	TTCCAAAAAA	AAAAAATAAA	ATTAC

**Figure 8** Nucleotide sequence of the *his-3 cog<sup>E</sup> lpl* region of linkage group I in the StLawrence wild type strain of *Neurospora crassa*. This differs from that in the Lindegren strain in many positions, summarised in figure 5. The coordinates of relevant features are given in the text. This sequence contains the weak recombinator *cog<sup>E</sup>* and also the remnant of a transposable element *Guest* within the replaceable sequence 3' of *his-3*. StLawrence strains carry *rec-2<sup>+</sup>* which prevents the initiation of recombination at *cog*.

1	ACCGGGAATC	GTAGCGGGCG	CTAAGGCCAA	GCCGCGGCAC	GGGTCACTGA	CCCAATGCAG
61	CGCATTCGGT	CAGCAACTGA	AGTGGATGTA	CAAGTACATA	GTAAGTATC	GCAACTGGAG
121	ATCACTCGCA	CCGTGCCGCA	GAACAAGGGC	GACGAGCCTC	AGGGCAGTTT	AGCCTGCCGT
181	AACAGCACAG	ACCATAGCTT	ATTTTCACCT	GGGCGGGCGG	GCGACGGCGG	CACTGACATC
241	GGCAAGGCGG	CATCAAGCAA	CCCCTCTGTT	GCTTGCCAGC	TGCCGGCCAA	CGTCAGCGGT
301	ACAAGGAGAA	ATCTGGAAGG	AAAGACTTCT	GGCACCGACA	GGATGGCAGC	CGGGAAAGT
361	TCCCAATGCA	TGAGATGAGG	GGCATTGTGA	TTGCCCTCCG	TCACCCAGTG	CGAACCCCAA
421	CCCCACCATA	GCGTCTGTCT	ATACATGGAG	CGCGAAGTCG	AGAAACCTGT	AATTCTTGGT
481	AACTTTCAGG	TACACAGTAC	GTACTGATCC	TGGTATCAAA	CCTTGCCTGC	CGAGTTTTTCG
541	ACGGAAAGAG	GTGTGAATTG	TGAAAGAGTC	ATACCAAATC	ACCCGATTTT	CATAAAGCCC
601	GAGTCTTTTC	TGTACATAAG	CGACACTCGA	AGCGGGCCTC	ATCTTCATAG	CCTGATAGCT
661	TGTAATACTC	CATCCTCGTA	TCTCACTTGA	CCTTGAGTTC	AACCCACAGT	CAAACCTTCAC
721	CCGACACATC	GACGATTGGG	GGAACAGCAC	AATACCTGAA	AAGCGAGAAA	ACCAAACAGA
781	GGAAACACCC	ATGGAGACAA	CACTTCCCCT	CCCCTTCCTC	GTCGGTGTCA	GTGTTCTCTC
841	CGGACTGAAT	GACATCAAGG	AGGGCCTCAG	CCGGGAGGAA	GTCTCGTGTC	TTGGCTGCGT
901	CTTCTTCGAG	GTCAAGCCCC	AGACCCTTGA	AGAAATCCTG	CGATTCTCTA	AGCGTCACAA
961	TGTCGAATTT	GAGCCCTACT	TGATGTAAC	AGCCCTCGAG	TCTATCGATG	ATATTATCAC
1021	TCTTCTGGAC	GCCGGCGCCC	CCAAGGTGTT	TGTCAAGACC	GAGCAGTTGG	CCGACCTCTC
1081	CGCATATGGC	TCCCGCGTTG	CCCCATTGT	CACTGGAAGC	AGCGCTGCTT	TGCTTTCCTC
1141	CGCCACCGAG	AGCGGCCTTT	TGCTCTCCGG	CTTCGATCAG	ACTGCCTCCG	AGGCTGCACA
1201	GTTTCTGGAG	GAGGCCAGAG	ACAAGAAAAT	TACCCCTTTC	TTCATCAAGC	CCGTTCTTGG
1261	GGCCGATCTC	GAACAGTTCA	TCCAGGTTCG	CGCCAAGGCT	AACGCCATCC	CCATCCTGCC
1321	ATCCACTGGC	TTGACAACAA	AGAAGGACGA	GGCCGGCAAG	CTTGCCATCT	CCACCATCCT
1381	CTCGAGCGTC	TGGAAGTCTG	ACCGTCCCGA	TGGTCTTCTC	CCCACCGTTG	TCGTTGATGA
1441	GCACGACACT	GCTCTGGGTC	TGGTCTACAG	CAGTGCCGAG	AGTGTGAACG	AGGCCCTCAG
1501	GACACAGACT	GGTGTCTATC	AGAGCCGGAA	GCGCGGTCTC	TGGTACAAGG	GTGCTACTTC
1561	CGGAGACACT	CAGGAGCTCG	TCCGCATCTC	GCTTGACTGC	GATAACGATG	CTCTCAAGTT
1621	TGTCGTGAAG	CAGAAGGGTC	GTTTCTGCCA	CCTCGATCAG	TCCGGCTGCT	TTGGTCAGCT
1681	CAAAGGCCCTT	CCCAAGCTCG	AGCAGACTTT	GATTTGAGG	AAACAGTCTG	CCCCCGAGGG
1741	CTCCTACACT	GCCCCTCTCT	TCTCCGATGA	GAAGCTAGTC	CGGGCCAAGA	TCATGGAGGA
1801	GGCTGAGGAG	CTCTGCACCG	CTCAGACCCC	CCAGGAAATC	GCCTTTGAGG	CTGCCGATCT
1861	CTTCTACTTT	GCTCTTACCA	GGGCCGTTGC	TGCCGGCGTT	ACTCTTGCCG	ATATCGAAAG
1921	GAGCCTTGAC	GCCAAGAGCT	GGAAGGTCAA	GCGCAGGACT	GGAGATGCTA	AGGGTAAGTG
1981	GGCTGAGAAG	GAGGGCATCA	AGCCTGCGGC	GTCCGCTCTC	GCTGCCACTT	CGGCCCTGT
2041	CACCAAGGAG	GCCGCCCAGG	AGACCAACCC	TGAGAAGATC	ACCATGAGAC	GTTTCGACGC
2101	CTCCAAGGTC	TCTACCGAGG	AGCTCGATGC	TGCTCTCAAG	CGTCCTGCGC	AAAAGTCGTC
2161	CGATGCCATC	TACAAGATCA	TTGTCCCCAT	CATCGAGGAC	GTCCGCAAGA	ACGGCGACAA
2221	GGCTGTTCTG	TCGTACACTC	ACAAGTTCTG	GAAGGCTACC	TCTCTTACTA	GCCCCGTCTT
2281	GAAGGCGCCC	TTCCCCAAGG	AGCTTATGCA	GCTCCCTGAG	GAGACCATTG	CTGCCATCGA
2341	CGTGTCCTTC	GAGAACATCC	GCAAGTTCCA	GCCCGCCCAG	AAGGAGGAGA	AGCCCTCCA
2401	GGTCGAGACC	ATGCCCGGTG	TTGTCTGCAG	CCGTTTCTCT	CGTCCCATCG	AGGCCGTGCG
2461	CTGCTACATC	CCCGGCGGTA	CCGCCGTTCT	CCCCAGCACT	GCCCTTATGC	TGGGTGTTCC
2521	CGCCATGGTC	GCCGGCTGCA	ACAAGATTGT	GTTCGCCTCT	CCTCCCCGCG	CCGACGGAAC
2581	CATCACTCCC	GAGATTGTCC	ACGTCGCTCA	CAAGGTTGGG	GCCGAGTCCA	TCGTGCTTGC
2641	CGGCGGTGTC	CAGGCCGTAG	CTGCCATGGC	CTACGGCACC	GAGAGCATCA	CCAAGGTCTG
2701	CAAGATTCTC	GGCCCCGGTA	ACGAGTTTCG	CACTGCTGCC	AAGATGTTTC	TCAGCAACGA
2761	CACCAACGCT	GCCGTTGGTA	TTGACATGCC	CGCTGGCCCC	TCCGAGGTGC	TGGTCATCGC
2821	TGACAAGGAC	GCCAACCCCG	CGTTCGTTGC	CTCGGATCTC	CTGTCCCAGG	CTGAGCACGG
2881	CGTTGACAGT	CAGGTTCATCC	TGATCGCTAT	TGACCTCGAC	GAGGAGCATC	TTCAGGCTAT
2941	TGAGGACGAG	GTTCAACGTC	AGGCTACGGA	GCTTCTCTCG	GTCCAGATTG	TCCGTGGCTC
3001	CATCGCCAC	TGCATCACCG	TGCAGGTCAA	GACCGTCGAG	GAGGCCATGG	AGCTCAGCAA
3061	CAAGTACGCT	CCTGAGCACT	TGATCCTCCA	GATCAAGGAG	GCCGAGAAGG	CTGTGATCT
3121	TGTCATGAAC	GCCGGTAGTG	TCTTCATTGG	CGCCTGGACT	CCTGAGTCCG	TTGGCGATTA
3181	CTCTGCTGGT	GTTAACCCT	CGCTGCGTAA	GTTACATATC	ATAAATAGCC	CCGCTTCACA
3241	GATTCTTCTG	CTAACGTCAA	GACACATAGC	TACCTATGGC	TTTGGCAAGC	AGTACTCTGG

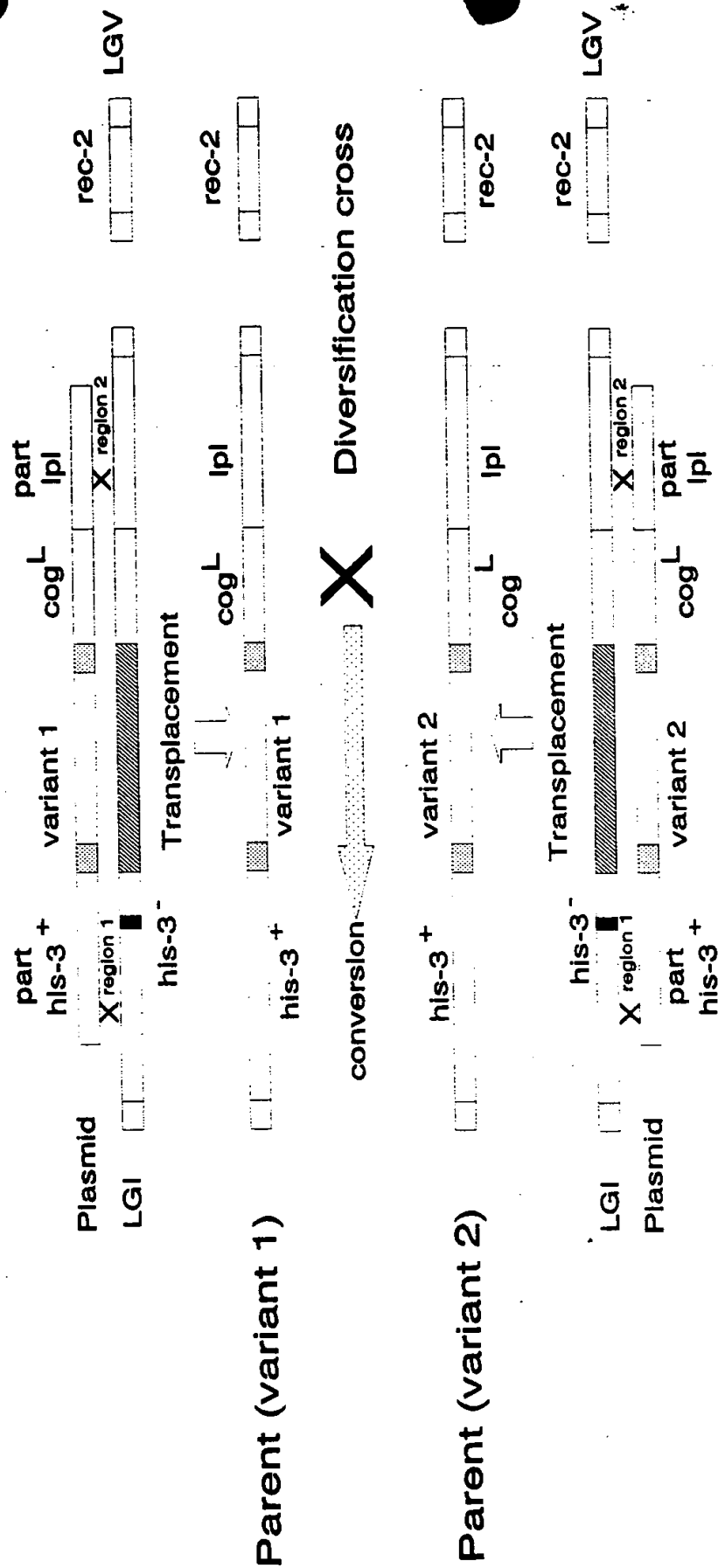
Figure 8 continued

3301	CGTCAATTTTC	GCCTCGTTTCG	TCAAGCACAT	TACCAGCTCC	AACTTGACTG	CCGAGGGTCT
3361	CAAAAACGTC	GGCCAGGCTG	TCATGCAGTT	GGCTAAGGTT	GAGGAGCTCG	AGGCTCACAG
3421	AAGGGCGGTC	AGCATCCGTC	TTGAGCACAT	GAGCAAGAGC	AACTAAACGG	AAATTCTTTT
3481	CGAAGTAGCA	AAAAAAAAAA	AAAAAAAAACAA	GAACAAAAGG	ATGTAGTGGG	TTGATGTATA
3541	TCTGGGTCAT	TTTGGGCACA	TAGAGTAATG	ATAACGAGTT	TTGGACATTG	TACTGTCTCTG
3601	TACAGGCTGA	AGATCAGTAC	ATGAATCTGT	TGGTAAGTGT	GGAGACCCAA	ACGTCCCTTG
3661	AGTTTTTCTC	CCTATTCCAG	AGGTGCTCGT	CCCTGGGTGT	TTATTTTCAT	TATTACATCA
3721	ACCTTTTTTT	TTTTTTTTTT	TTTTTCAGAT	CATGCGTACA	TGAACGGGGG	AAGCACAGAC
3781	GATCGAAACG	TGGATGTCAC	AATGTCGCTG	CAGTGATGCT	GCATTGCATG	AAGCGCCCAT
3841	CTCAATATAC	TTGCAGTCTT	GCACGTTGCA	TGTGAAC TTC	CCAAACAACC	GAATAAAAGA
3901	CGGCGAAAAA	TGAAGATAAA	AAAAAACCAT	AAAAAAATC	AGAGGGAGTG	TGGGAAATGG
3961	TGTCTTTTAG	CATTACAGACC	CCATAGCCGT	GCACGCCCGG	GTACAGACAG	GTTCATCGAT
4021	GTTGACATTG	ACTGGGACAC	CAGGTCATAT	TATTTTATCT	CCTGTCTCTT	ACCATACATC
4081	GGGACATCGG	ACATCTTGCT	GTACCCCCCA	CACCCACAAA	GCCTTATAAA	AGCGCCACAC
4141	CCGAGGAGGT	TCGGTCCGGC	CCACGAACTC	TGTGCCTCCC	TGCCTGTTTA	CAGGGACCGA
4201	ACGCTGGAGA	ATCTTACTAG	TTTCTTGACA	TCCGGCCTAC	CCGAGCAGGA	AAAGGGACAG
4261	CTCATAGGCG	AGGAGGGATT	TGAAGATGGG	AACATTTTGG	GTGATTCGAG	AGGAGGAAC
4321	AGGTACTGCA	TCATGATAGT	TCGGGGCAGC	ATCTTGGCTG	GGACATTGTT	AATACCTCGA
4381	TATGATGAAG	TAGGAGGGAG	TTTTTGCGTG	TCTTGCCGAA	GTCCAGAGAT	CTGTTTTATT
4441	TTATTTTTTA	TGGATGTAGT	GTATCAACAC	CCAAGATTCTG	GAGAATAGTA	CTAGGATTCTG
4501	CATTACAAAG	TGGAAGTCTT	GAGAATCGTT	GTATATCCTT	GTCTTCTCTG	GAATGTTAAC
4561	AATCCTACAG	CSAGCGAGCG	AGCGGTCCGA	TGCGCTGATC	TGATAGGCGC	AATATCCGGC
4621	CGCTTTCTCC	GGTCGTGTAG	TCTAAGCTCT	GTGGGCATAG	TACACTAAAA	AAACCCTTGC
4681	ATTTTCATGAT	CTGCCTGCTA	TTCATTCCGA	GCTATTTTCTG	TGGTCACATT	TCGAGGAAGA
4741	AAGAAAGCAA	CTAAGATTCA	CAGCCATCCA	TCCATCCATA	TGGAAGAATA	ATCCATTCCC
4801	ATGTTCCCTC	CCCCCCTACT	TGTATGTGAC	CACACGCTGC	TGTCAGAATG	CCAACGGTCT
4861	CAGGTACCCT	CGTCCGACTG	TTTGCCATGG	AGTTACATAC	ACTACTAGTG	TAGCCCCGGG
4921	CCAAGTACC	CCGTCAAATC	TATACATATC	TATAACGGGT	TTTACGGGTT	TCGTTTCGCTG
4981	TCAATCAAGT	TTGAAACATC	ACTGGGGCCG	TTGGACGGTG	TATTAGACCA	TTGGCTCCCT
5041	CAGCTGTTTG	GCGGCTGGGC	GGCTGGGTCA	ACGGGCAATA	ACGGGACTCG	AGAGGGACGA
5101	GGAGAGTCGG	TTGGCTGGCT	GCAATACAAG	CGTTCCCAAC	TAACCAACGA	GTCCCGTTTT
5161	CCATTTGTGT	GCCTAACCAT	CATCTAGGGA	TGTCAGGGT	TGGCCGGATC	AGGGTATGTT
5221	TGGTTGACTG	TTGTCATGTC	TGATTGGGTA	CATATTATGG	TAGGTGTCTC	GAGAACAGTA
5281	GAGTACTCGG	GCCTAGCGTT	TGGATGATTA	CGCGAGATAT	GAGTTGTGGG	CCGCCATGCA
5341	GTTGCTTGTC	CATAAGCAGA	AGTTGCTTTG	GGATATATTT	CTCGTCTTTC	AAAGGTCACG
5401	AGGTCCCTGGG	ACGAACGGCA	TCGCCATCCA	AAGGGTTGAA	CATGAGAAAC	CTGAATGGCC
5461	TTTGCGTTGA	AATACAAAAA	GTCAAGAACA	AAATCGCTTG	AGGATAGGGA	CGTGGGAAGCA
5521	AGCAAATATG	GTAAGAGAGG	TATACATCAA	CCCTGGTTCA	ATTGTTAGCG	TGGTTCTTCC
5581	TCCACGTCCT	CGTTCATGAC	GGTTAACAGT	ACCAGGCTAA	CAATTAAACC	AGGGTTGATG
5641	TGTACTGATA	TGTAGGTGCT	CAGCAAATCG	CCAATTTCTT	TGGCCCCAAG	TAGAGCGTTG
5701	CTGTCAGTGC	TGCTCGTGTC	AGCCTTGGTA	TGGGAACCTA	AACTGCTAAC	ACAGCGCAAG
5761	TGCGCATGTA	AAGATATTGT	GGGAGGATCT	GTATGGATGG	ATGAGATTAC	TGCTTGGTGT
5821	TGGTTGCGAG	GCACTGCGGC	TGTTAGGCTT	TGCTGTGCCC	CGTTCGACGA	AGAAATACCG
5881	GGAACATATA	ATTGGATACC	TAGACTTACT	GCCTATGGGA	GGTATCTACC	GACGTAGCCG
5941	ACGGATTCTA	GCAACATCCC	GACTTTGCTT	GTAGTGTACT	ATGATAGCAG	CACAGTGTG
6001	CTCCTTGTGA	GAATGGGCTC	TTTTTTTTTT	TCCCCCTTCC	CTAGGGCGTT	GACTGGACTT
6061	GCTCTATTGT	TCCCAAGGTA	GGTGCCCGTC	ATCGATTTTC	CCAAGTCTCC	CGCCAGATTG
6121	TCGTCATAGT	GTCATGATGA	CCTCGGTCGC	TGGGGCTGCG	TGGTTACGGG	GAGCTGGGAC
6181	CGCTAGGCCT	CAGTGGTTGT	GCCATTGAGC	GTGGGTGTGT	GGAGTAGCGG	TAGAGGCGCT
6241	TGGAAGTTGT	GCTAGCGGAA	ACCCTGGAAT	ATCTTCTACC	CTCGATTCTT	TCTCGGGCTG
6301	CCCATGTGCT	GAGGTGATGC	CGGGGATCTG	GCGCCAATCA	TCCATTGAGG	TTCCCGCAGC
6361	TTCCCGGTGC	CGCGCGCGGG	CGCAGTTGCT	CACAGGACAC	ACCTAGACGC	AGGGGCACAG
6421	GGGCACCGTT	TGGTGTGCAA	CTGGGTACCT	AGCTGTAGCA	AGCACTCCAC	CGTCTGTGCA
6481	ATCCCCCAAT	CCACGGCAGG	AACTTCGCAC	CGCCGCGGCA	CCGAGTGAGC	GAATCCATCC
6541	GCATTGGATC	CCAATTCTTG	CCCTTGCCAT	CCTTCTTTCT	TCCCACTTGG	CGCAACCAAC
6601	ACTTCCCTTG	GTCTGGGTAC	TCGTGTTGAT	CTTCACTCTC	TTTTTTTCTT	GGGCGACCGA
6661	CTTTTTATAT	CCGTCCCTTG	TTCCCCCTGG	CCGTTGTCTG	TCTTTCTACA	ACTACCTTCC
6721	GTTTATTATC	CCCTTTCTTG	GTTCGGTCGA	GGACCCAAAA	ACAGAACAAT	TCCGGCTCTT
6781	CCAGGTGGCT	TGGGTGCGAC	TGTTTAGCTC	TTGACCACTA	GCCGCTTACC	TCTCTTGAT
6841	GTTTTTATTT	GGATATCATT	AAACTACTCT	TTCTTGAAAC	GGCAGACGAA	CGGAACAGTT
6901	CCTACGGTAT	ATTAGCGATA	TACGTTGTAC	TGATATTCTG	AGCAAGAAGA	GGCAAATTAT
6961	CAATTATGCA	TCTCCCTTCG	TCGCTGCTCA	TCGCAGCTCC	CTTGCTCGCC	AATGTATCGG
7021	CCGAACCCAT	TAGGATACCC	CAACGCGATG	TTCTCCGTGG	TATCAACATC	ACAGCAACTT
7081	GCCGTTTCAG	CACCTACCGA	TTCGCCAGC	GGTGGATATG	CCCCTGCCGT	TGTAGACTGT
7141	CCCAAGACCA	AGCCGACGCT	CCGGAAGGCC	GTGGATTTGT	CGAACGAGGA	GAAGAAGTGG
7201	TTGTGATATC	GGAGGAAGAA	CACCATCCAG	CCCATGAGGG	ACCTCCTGAA	GAGGGCCAAC

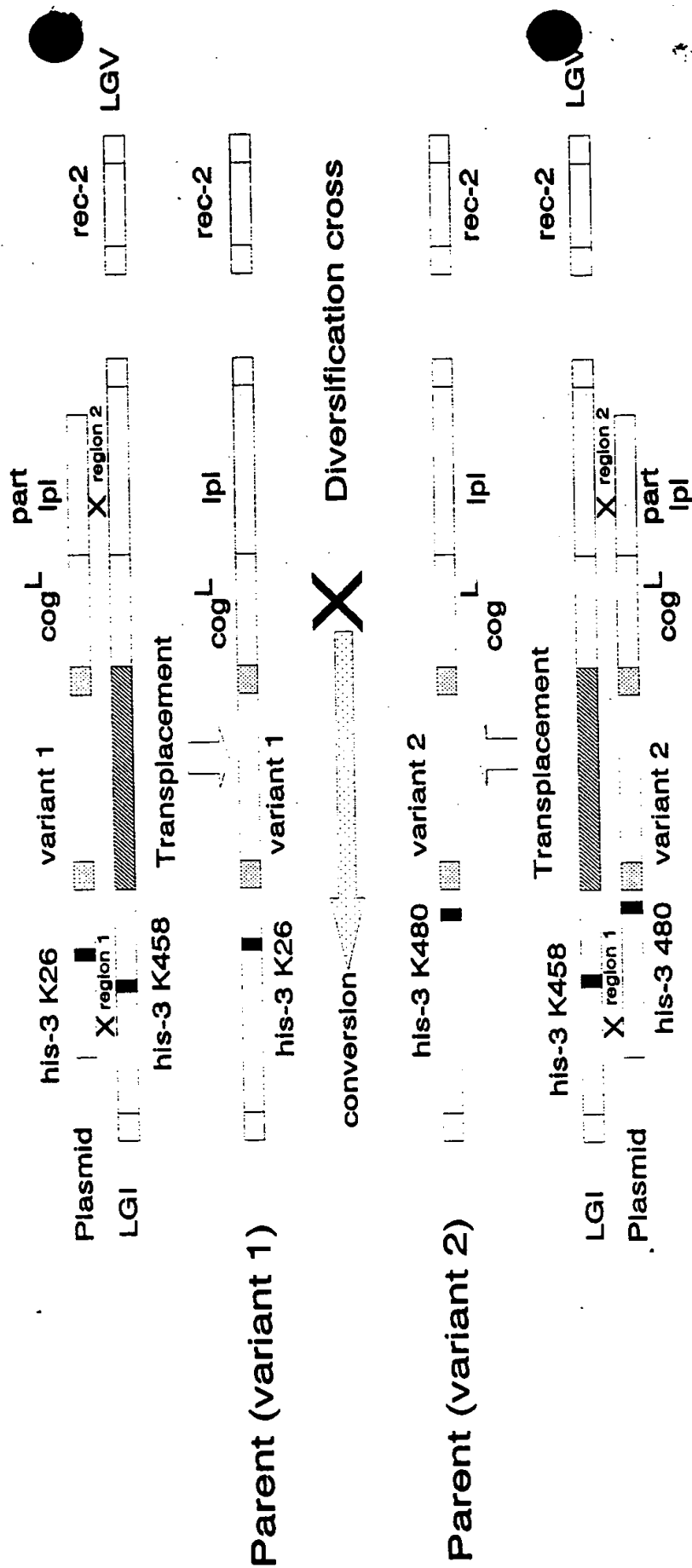
Figure 8 continued

7261	ATCACTGGGT	TCGATTCCGA	GACATTTATG	AATGAGGCCG	CCAACAACAT	CTCGCAACTG
7321	CCCAATGTCG	CCATTGCCAT	TTCAGGAGGC	GGCTATCGTG	CCCTCATGAA	CGGCGCCGGC
7381	TTCGTTGCTG	CTGCGGATAA	CCGAAITCAA	AATACCACGG	GCGCAGGTGG	TATTGGAGGC
7441	TTGTTGCAGT	CCAGCACATA	TTTGTATGTA	AAGTGGTTCT	TCTTATCTCG	TTTTCGAGTG
7501	TCAACTGCGC	CAGTTCAGAG	TTGGGCGGCT	GTGGACGACC	TTGCTGGTGA	ACATGTCTTG
7561	GACTCCATGC	CCCTTCTTCG	TTTCCTCAAA	TCAAGAAGTC	GAGGACCGTG	ACCGTAAATC
7621	GCTAACGCAA	CTCTAGGGCC	GGACTTTCTG	GTGGTGGCTG	GCTTGTCCGC	AGTTTGTTCT
7681	CCAACAACCT	CAGCAGCATT	GAGACCCTGC	TGAGCGAGAA	CAAAGTCTGG	GACTTTGAGA
7741	ACTCCATCTT	TAAAGGGCCC	AAGGAGGCTG	GCCTTAGTAC	TGTCAACCGC	ATTCAAGTACT
7801	GGTCCGAAGT	GGCAAAGGAA	GTTGCCAAGA	AGAAGGATGC	TGGCTTCGAG	ACAAGTATAA
7861	CAGACTACTG	GGGCGGAGCA	TTGAGTTACC	AAC TGATCGG	AGCCGATATG	GGCGGCCCGG
7921	CTTACACCTT	CTCCAGCATT	GCCCAGACCG	ACAAC TTCCA	GAAGGCCGAA	ACGCCGTTCC
7981	CTATTCTGGT	AGCTGACGGC	CGCGCGCCTG	GAGACACCAT	CATCTCCCTC	AATGCTACCA
8041	ACTACGAGTT	CAACCCGTTT	GAGACGGGTA	GCTGGGACCC	GACCGTCTAT	GGCTTTGCGC
8101	CGACCAAGTA	CCTCGGCGCC	AAC TT CAGCA	ACGGCGTGAT	CCCATCGGGA	GGCAAGTGCG
8161	TTGAGGGTCT	CGACCAAGCC	GGCTTCGTCA	TGGGCACCAG	CAGCAGCTC	TTCAACCACT
8221	TCCTTTTGGC	CAACATCTCC	AGCTACGACG	GTGTTGCCCG	ACGTGCTCAT	CGAAGCCGTG
8281	ACTTCTGTCC	TCAAGGAAAT	CGGCGCCAAG	AGGACGACGT	CTCCCAAATC	ATCCCTAATC
8341	CGTTCCTGGA	CTGGAACAAC	CGGACCAACC	CCAACGCCGA	CACGCTCGAG	CTCGACCTGG
8401	TCGACGGCGG	CGAAGATCTG	CAGAATATTC	CGCTCAACCC	GCTCACCCTA	CCCGTGCGCG
8461	CCGTGGACGT	CATCTTCGCT	GTCGACTCGT	CCGCCGACGT	GACAAACTGG	CCCAATGGCA
8521	CGGCCCTGCG	AGCCACCTAC	GAGCGCACTT	TCGGCTCTAT	TTCCAACGGG	ACACTCTTCC
8581	CCTCGATCCC	CGACGACTGG	ACGTTTATAA	ACCTAGGCCT	CAACAACCGC	CCCTCTTTCT
8641	TCGGCTGCGA	TGTTAAGAAC	TTTACCTTGA	ACGCCAACCA	AAAGGTTCCC	CCCTTAATCG
8701	TCTATGTCCC	CAACGCGCCC	TATACCGCGC	TGAGCAACGT	GTCCACCTTC	GATCCGTCAT
8761	ACACCATGTC	TCAGCGCAAC	GACATCATCG	GCAACGGATG	GAAC TCAGCC	ACGCAGGGAA
8821	ACGGCACGCT	GGATTCCGAG	TGGCCCACTT	GCGTCGCCTG	CGCGGTTATC	AGCAGGAGCT
8881	TAGATCGGTT	GGGCAGGCAG	ACGCCAGCCG	CGTGCAAGAC	TTGCTTTGAG	AGGTATTGCT
8941	GGAAATGGCAC	AGTGAAC TCA	AAAGATAACAG	GGGTTTACAT	GCCTGAGTTC	AAGATTGCGG
9001	ATGCGCATGC	CCTGGACTCG	GGTGCTGTTG	CTATCGGAAA	GATGGTGAAT	GTCTGGTCGT
9061	CGGTTGTGGT	GGGAGTTGTG	GCGGCTACTT	TGTTGTTGTA	GGGGTAGGGG	AGACGTGATG
9121	ATATTCCAGT	CTGATGAAGT	TGAGACTGGA	CTGGAGATCG	CCAAGGATGC	GGAGGGAAAG
9181	GAATGCGTGG	TGTTAATGTC	ATGATGGATG	AAGGGTCATG	GATCATGGAA	CGACGGGGCG
9241	GGGATATTGG	ATGATGGATA	TACCACACTG	CATGCATGCT	CTATTGATAA	TATGCTTTGG
9301	CATTTACGTT	TAACAATCAA	TTGCTCCATC	CTGATGTTCT	ATCTTTCGAC	ACTGGATTGA
9361	TACTACTCCT	GTTGCTTCCC	TCTTGAAGTT	GGAAAGACTT	GAGGTTGGAA	GGACTTGAGG
9421	TTGTTTGTTC	TGAGGGAGGT	TATCGAAGTA	TCATCTGTGC	TGATGCCGAT	CGATAGACTG
9481	CCCTCTCTCT	CGAGGCAACG	AACGGTCGGA	TGAGCCTCTA	ATCATGATGC	TCAGTGCCAC
9541	AAAAAGGCTC	CAGCACAGCT	GCCCACACCT	TTTTTGCCCT	GTCGCTCCTT	CCTTTTTTTC
9601	CCCCCCTTTC	TTCTTTTCCA	TCTCATCCCG	TACCAGAGTG	CCCACCGGGT	ATATATATTA
9661	CCTCCTTGGC	CGTTCTCCTT	TGACCAATAA	ATCGCTTGGT	CGAGTGCGGT	AACCGTTTAC
9721	CGTCTACACT	TATCACTCAA	ACCAAACCAA	ACCATCGAAG	AAGTTACCTA	TCGGTTCGAG
9781	GGAAACGGTGA	TGTTCTTACG	TTCAAGTTAA	CCCAAAGAGC	GTTCCACATC	GTTGAACCGT
9841	CTCCTCCAGT	TCTTGGATCT	GTTTAACTTC	CGCAGCGACT	GAAGAAGTAA	TCACTTTTTT
9901	TTTTTTTGGT	TCCAAAAAAA	AAAAAAAATA	TTAC		

**Figure 9** Construction of the components of the sequence diversification cross: Parent (variant 1) and Parent (variant 2). For convenience, plasmid sequences are shown as linear. The cross hatched region in the chromosome is dispensable. Stippled sequences in the plasmid indicate the multiple cloning site for inserting foreign DNA. Crossovers in region 1 and region 2 insert the foreign sequence to be diversified into chromosome 1 of *Neurospora crassa* adjacent to the recombination hotspot *cog*. Parent (variant 2) containing a version of the foreign sequence with multiple differences from that in parent (variant 1) is similarly constructed. Parent (variant 1) and parent (variant 2) are crossed and conversion events (stippled arrow) initiated (X) at *cog<sup>L</sup>* recombine the sequence differences in variant 1 and variant 2 to form new combinations. Sequences are identical except for those that distinguish variant 1 and variant 2. *rec-2* on linkage group V permits *cog<sup>L</sup>* to be active. For simplicity, genes not directly related to the diversification are omitted. See text for further details.



**Figure 10** Construction of parent (variant 1) and parent (variant 2) enabling selection of progeny that have experienced conversion in the foreign DNA. Complementing pairs of *his-3* alleles are used to obtain parent (variant 1) and a different pair of complementing *his-3* alleles are used to obtain parent (variant 2) as explained in the text. Parent (variant 1) and parent (variant 2) are crossed and *his*<sup>+</sup> recombinants are selected. These must all have experienced conversion events affecting the foreign DNA since the events begin at *cog*<sup>L</sup>. The *his-3* alleles in parent (variant 1) and parent (variant 2) are non complementing to ensure that selection yields recombinants and not aneuploid progeny having two copies of all or part of linkage group I.



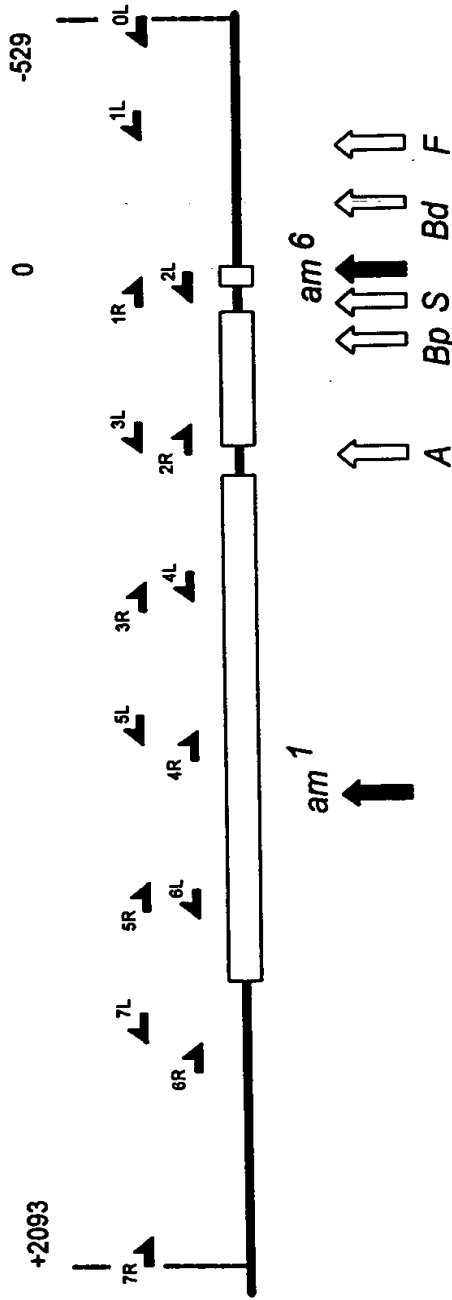


Fig 11

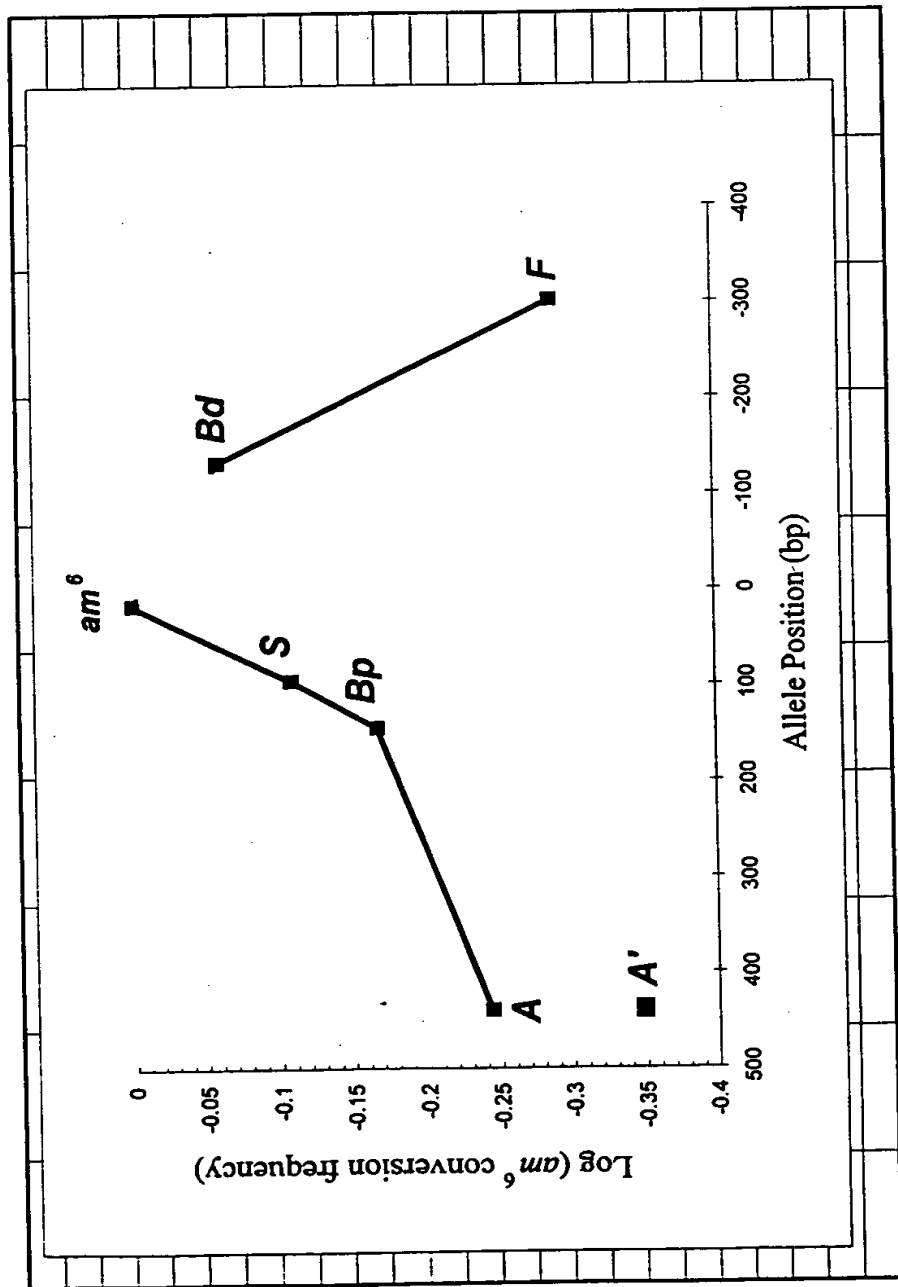
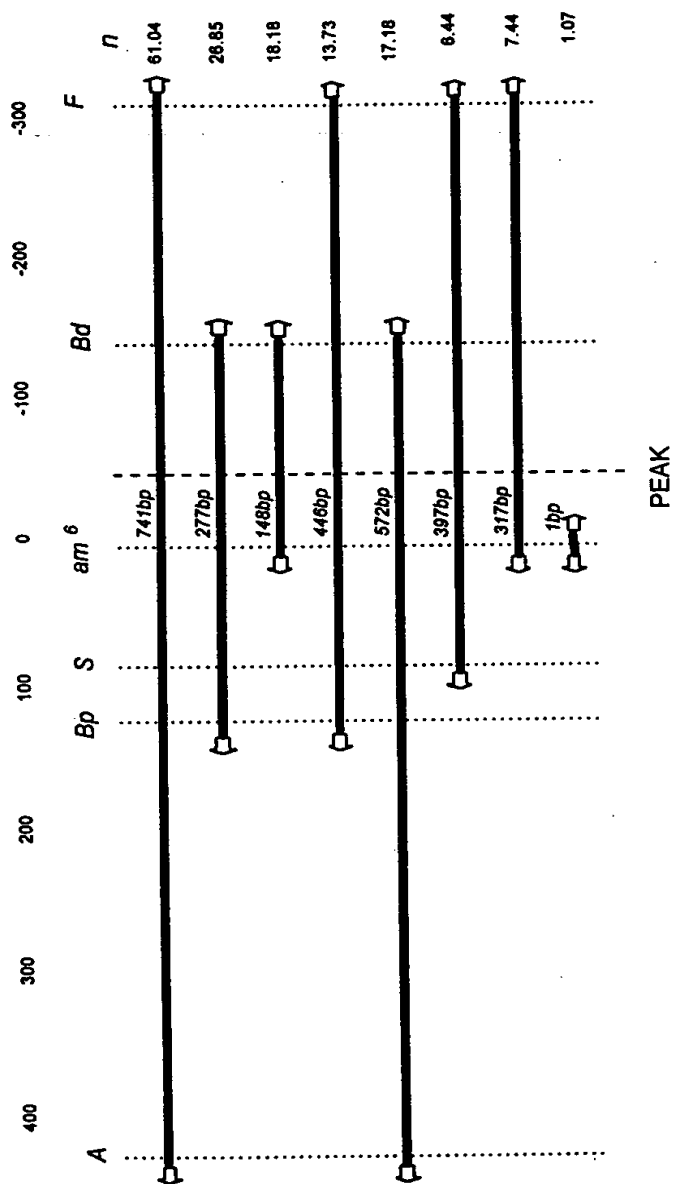
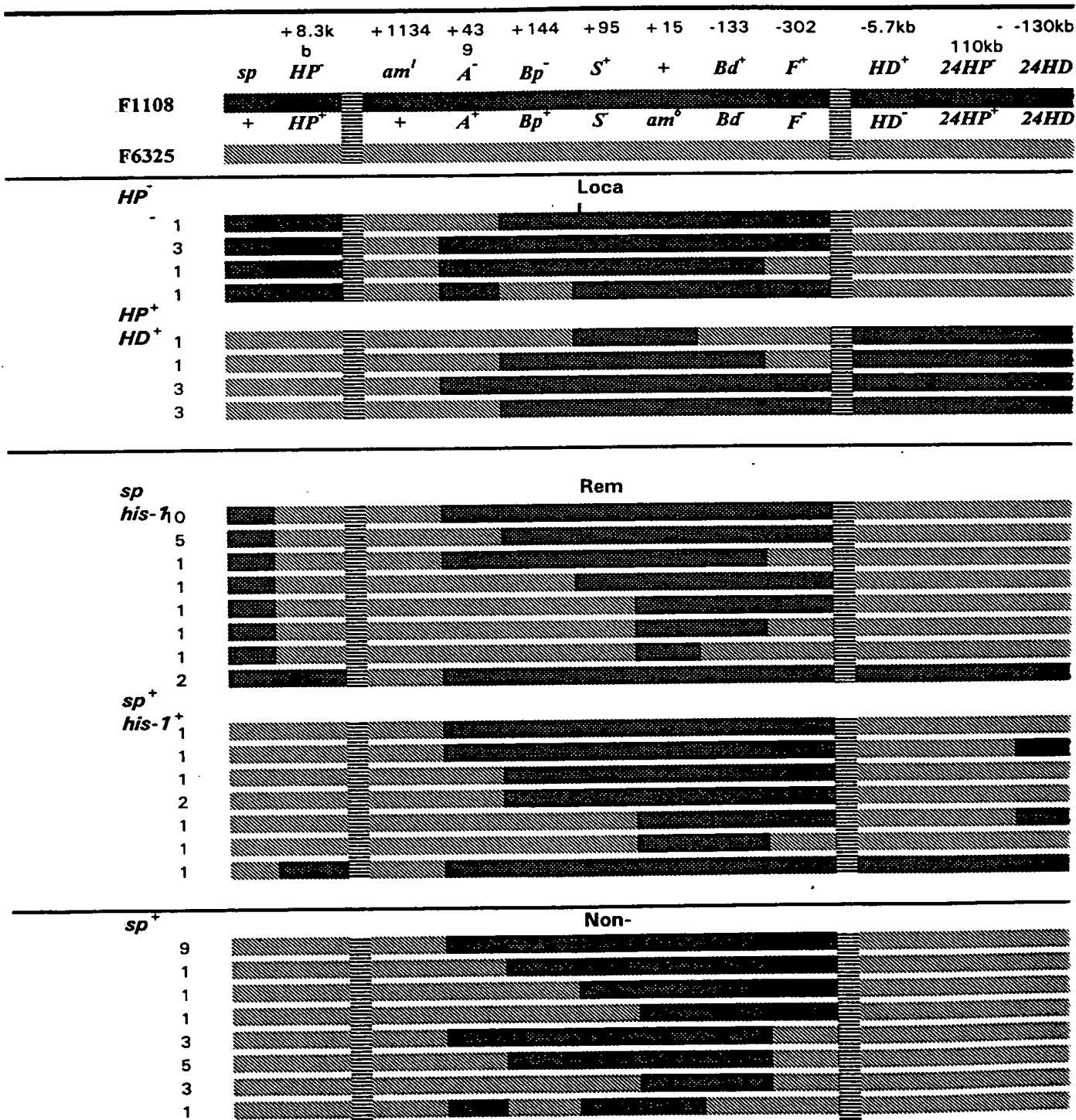


Fig 12



[illegible]

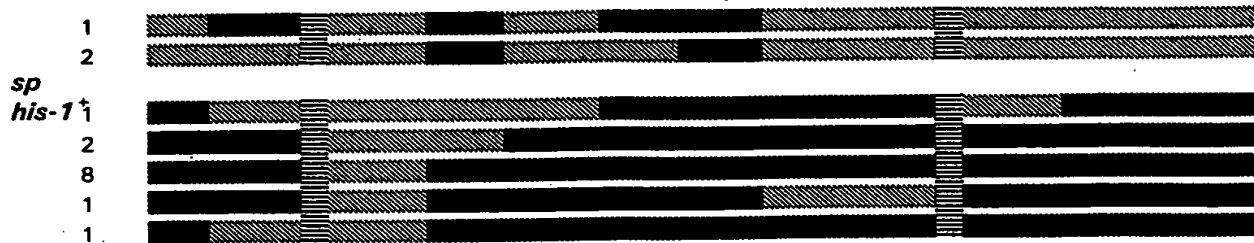


Fig. 14 (continued)

FIG. 1A

Methods for the diversification of DNA sequences and testing for superior variants

existing protocols: *Number of transfections needed to generate 1024 new variants: 1024*

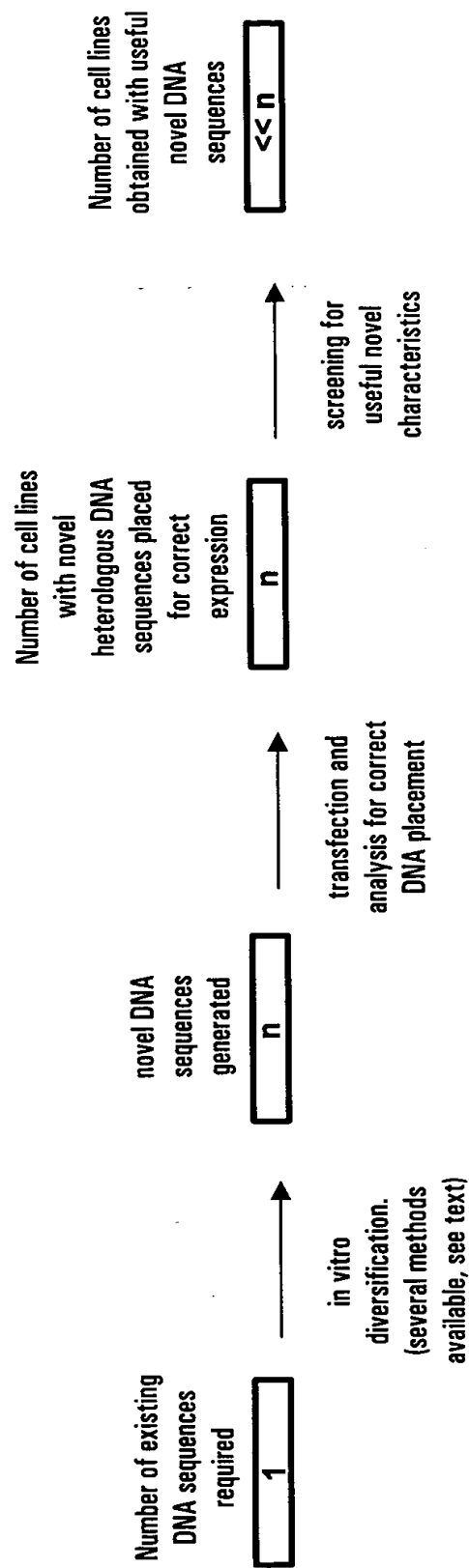


FIG. 1B

A protocol enabled by the present invention: *Number of transfections needed to generate 1024 new variants: 2*

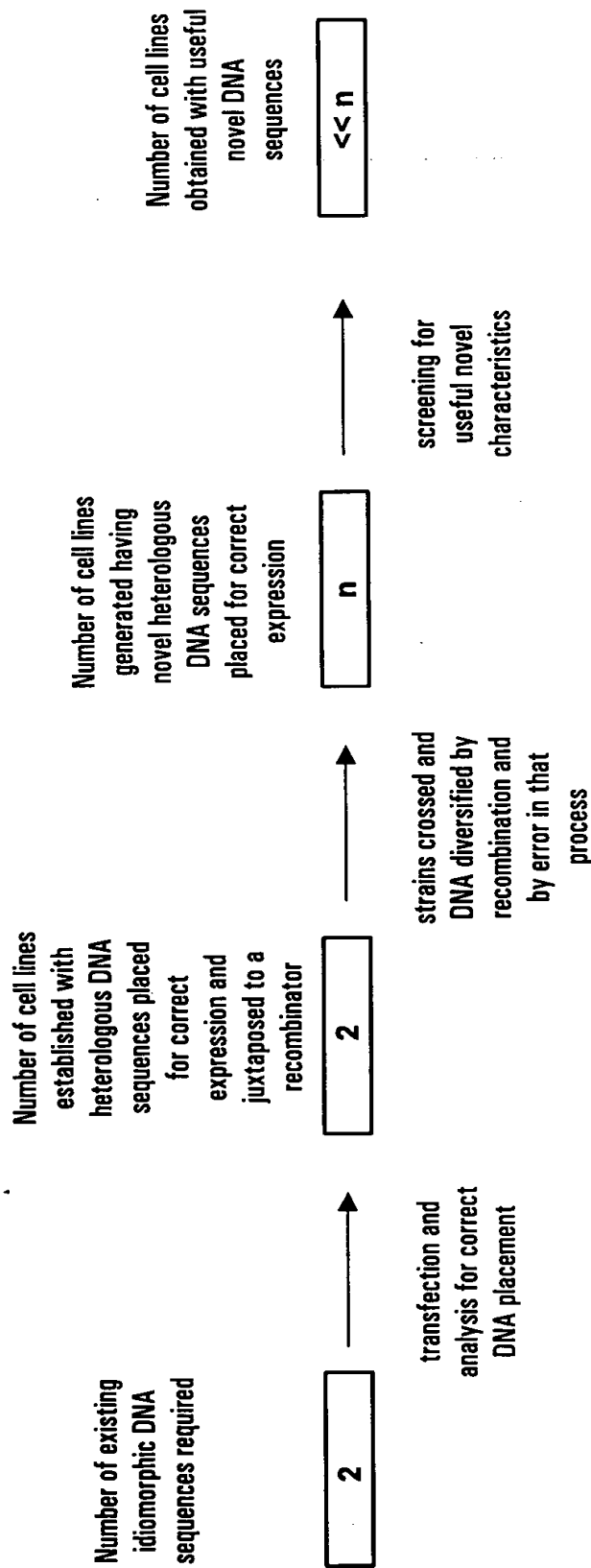


FIG. 2A

Methods for the diversification of DNA sequences coding subunits of heteropolymeric proteins and testing for superior variants.

The example given for immunoglobulins is for illustrative purposes only and is not intended to limit application of the present invention to this specific heteromeric protein. H = heavy chain genes, L = light chain genes

Existing protocol: *Number of transfections needed to generate 1024 new combinations:* 2048

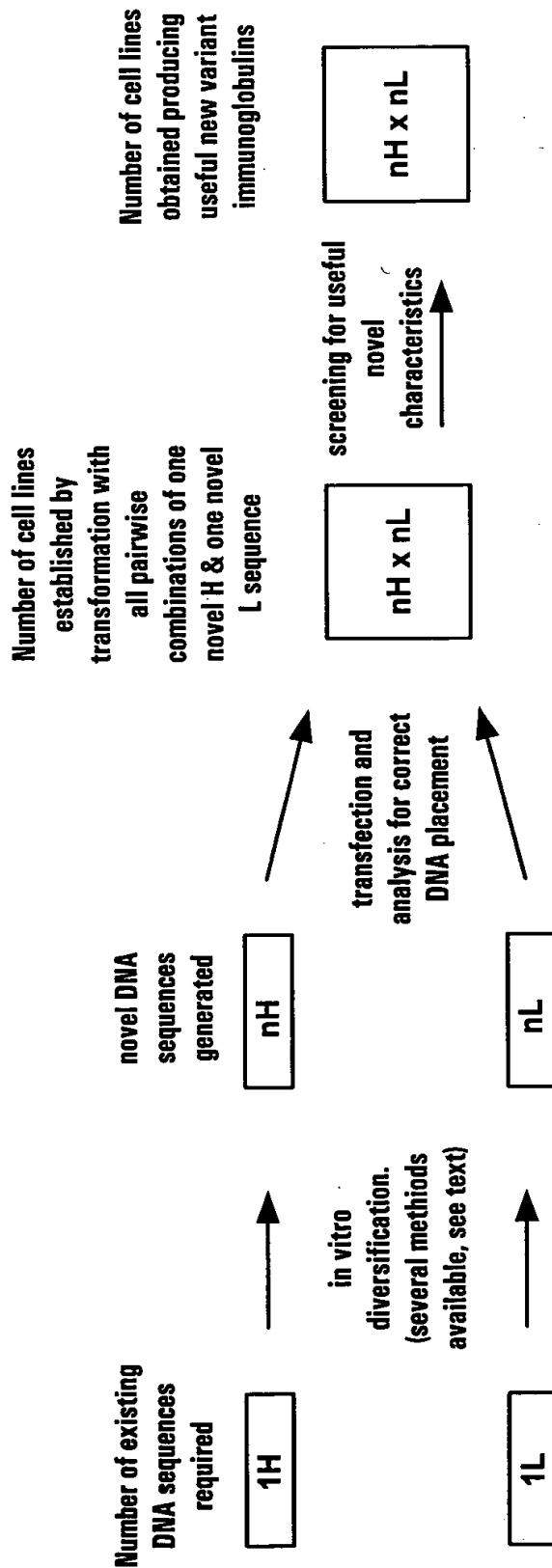


FIG. 2B

Existing protocol using the heterokaryon technology of US Patent Serial No. 5,643,745  
*Number of transfections needed to generate 1024 new combinations: 64*

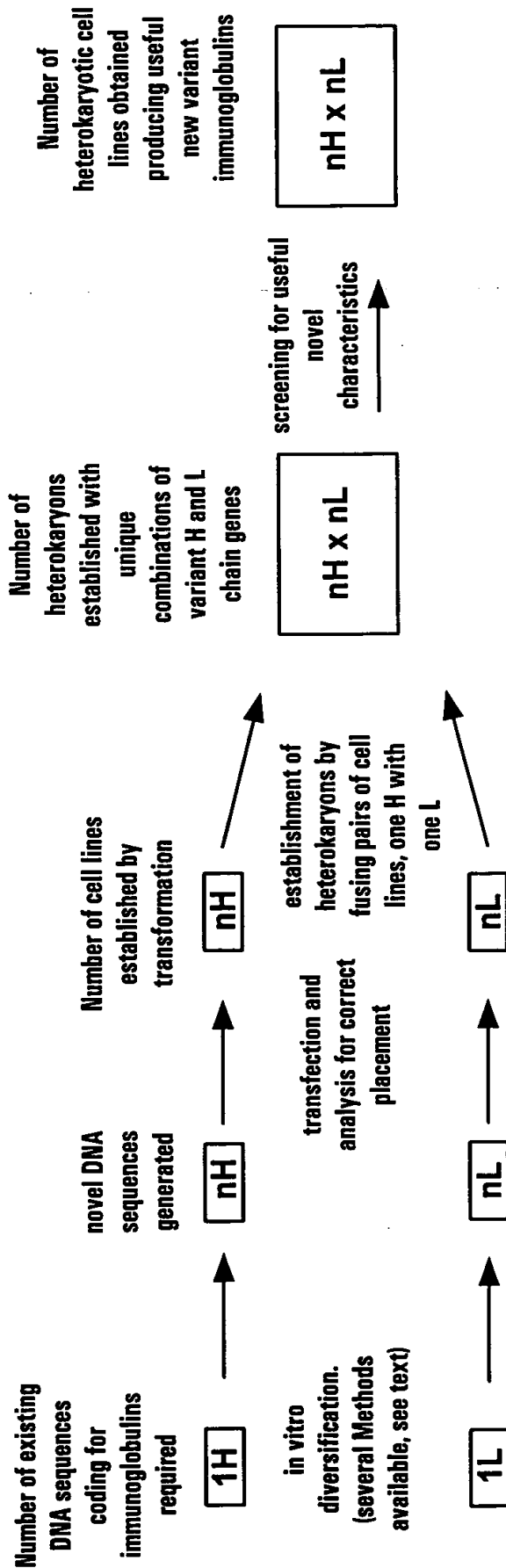
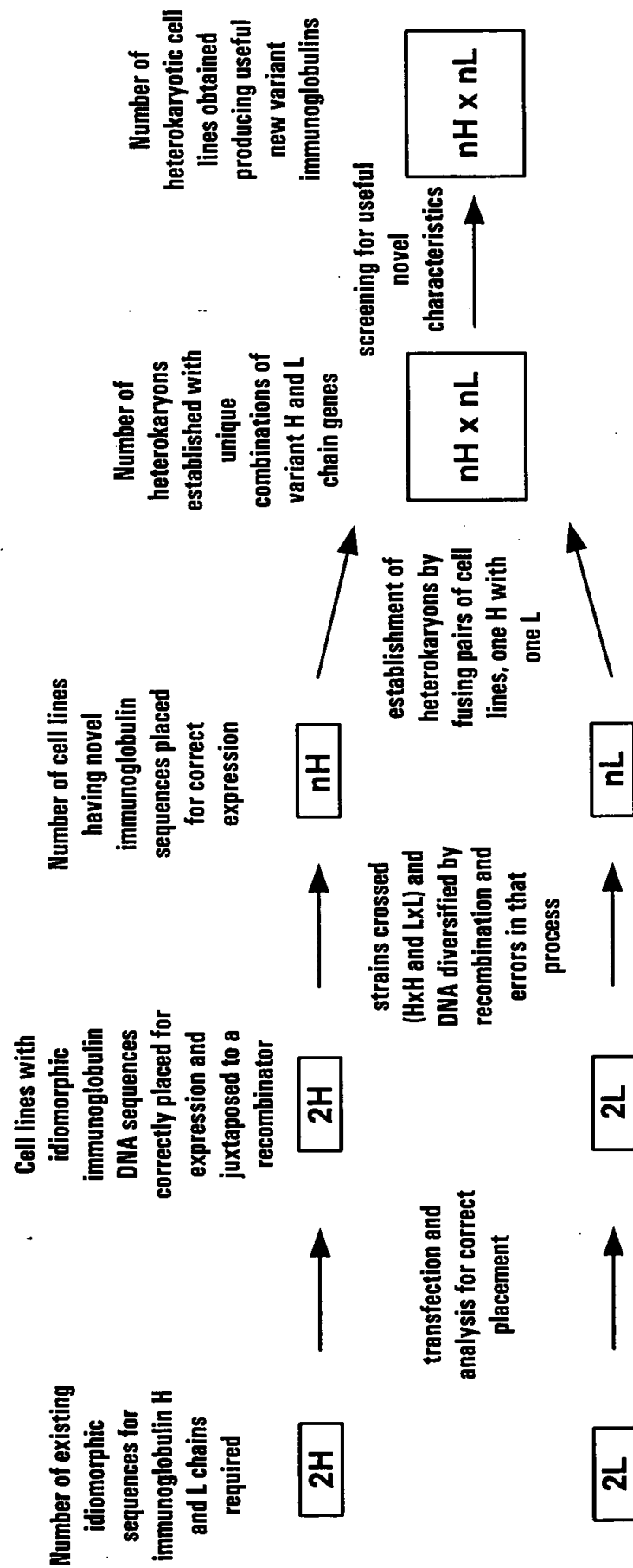


FIG. 2C

A protocol enabled by the present invention

Number of transfections needed to generate 1024 new combinations: 4



## Fig. 3

The modified double strand break repair model for meiotic recombination. After H Sun *et al* Cell **64**: 1155-1161, 1991

(a) A double strand break (DSB) is made in one DNA duplex. (b) A long 3' overhanging single strand tail is generated either side of the break by resection. (c) One 3' end invades a homologous duplex forming a D loop. (d) the D loop is enlarged by repair synthesis and anneals to the second 3' end (e) Repair synthesis occurs at the second 3' end and two intermolecular junctions (Holliday junctions) are formed. Resolution of the junctions by cutting inner and outer strands can give rise to non-crossover (f) and crossover (g) chromosomes. If there are base mismatches in the heteroduplex regions (duplex molecules with thick and thin lines) there will be gene conversion. If mismatch repair does not occur there will be post meiotic segregation of new sequence combinations.

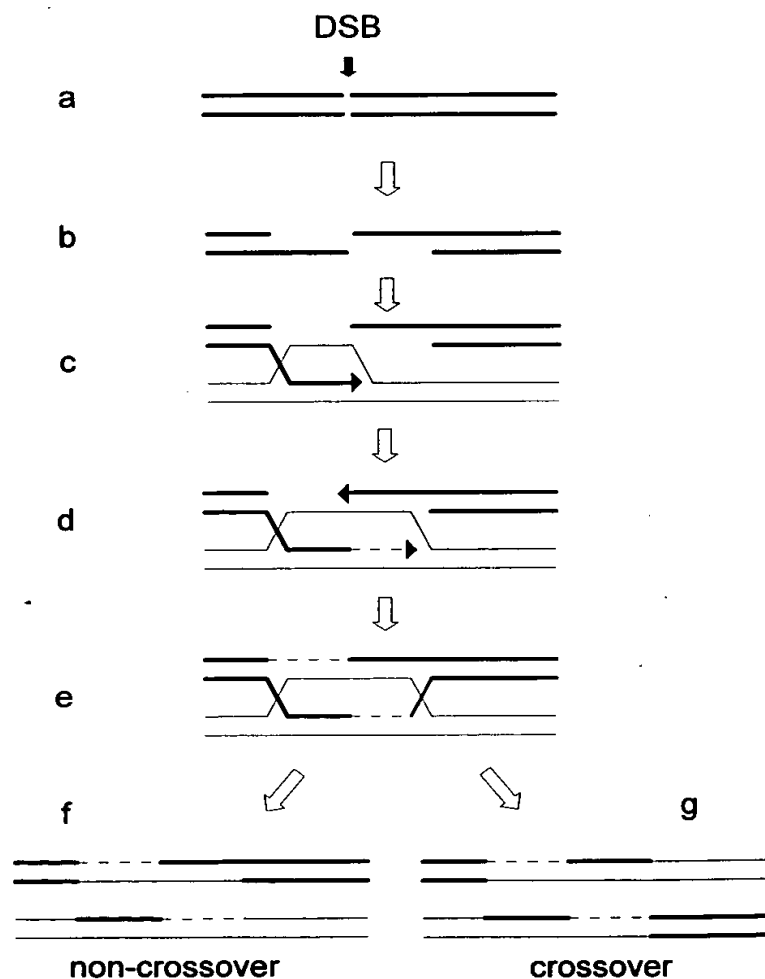
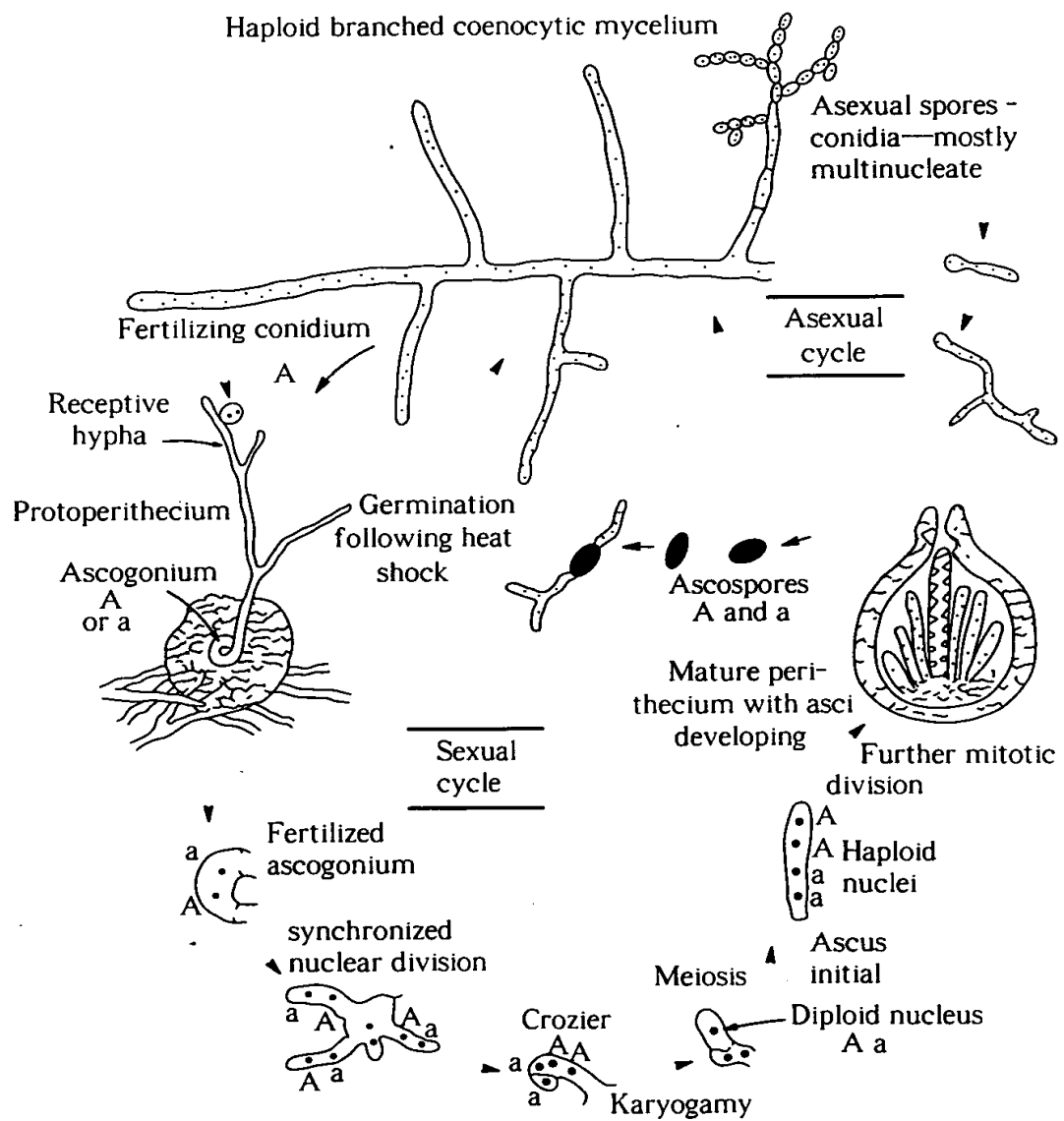


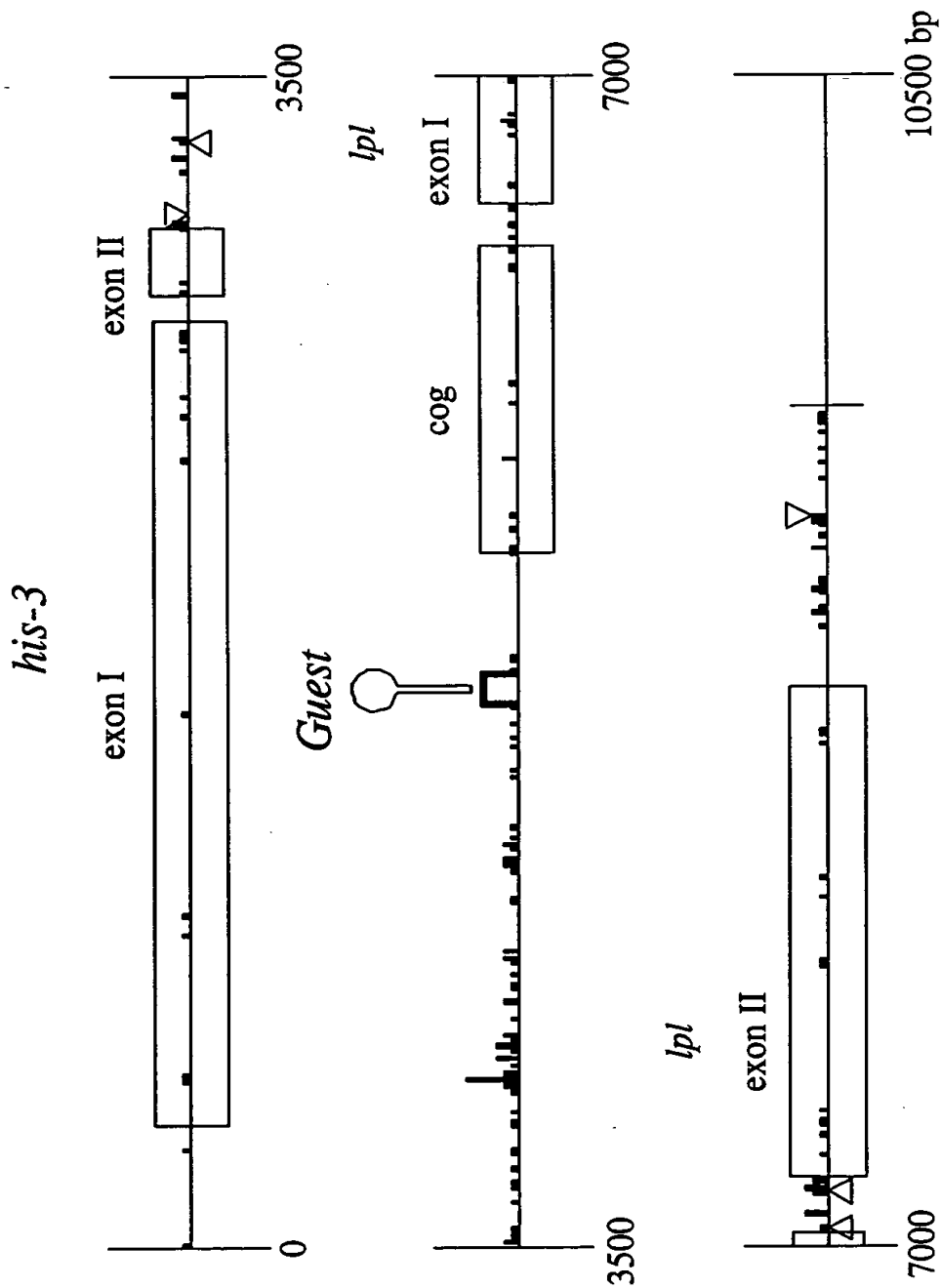
FIG. 4

Life cycle of *Neurospora crassa* after JRS Fincham (Genetics, Wright 1983). Microconidia having one nucleus are not shown but can be generated as described in the text. Perithecia and protoperithecia are shown in section.



**FIG. 5**

Map of the *his-3*, *cog*, *lpl* region of Linkage Group I of *Neurospora crassa*. Vertical bars, triangles and hairpins show the location of sequence differences that distinguish the St Lawrence and Lindegren wild type strains. The corresponding full DNA sequences are given in Fig. 7 and Fig. 8. Vertical slashes indicate one to seven base substitutions per 10 base pairs. Triangles indicate short sequence insertions and the hairpin a 101 base pair inverted repeat transposon fragment present in St Lawrence.





# FIG. 7

Nucleotide sequence of the *his-3 cog<sup>L</sup> lpl* region of linkage group I in the Lindegren wild type strain of *Neurospora crassa*. This differs from that in the StLawrence strain in many positions, summarised in figure 5. The coordinates of relevant features are given in the text. This sequence contains the high frequency recombinator *cog<sup>L</sup>* which is active providing the cross in which meiosis occurs is homozygous *rec-2*.

```

1  GATCGCAACT  GGAGATCACT  CGCACCGTGC  CGCAGAACAA  GGGCGACGAG  CCTCAGGGCA
61  GTTTAGCCTG  CCGTAACAGC  ACAGACCATA  GCTTATTTTC  ACCTGGGCGG  GCGGGCGACG
121 GCGGCACTGA  CATCGGCAAG  GCGGCATCAA  GCAACCCCTC  TGTGCTTGC  CAGCTGCCGG
181 CCAACGTCAG  CCGTACAAGG  AGAAATCTGG  AAGGAAAGAC  TTCTGGCACC  GACAGGATGG
241 CACGCGGGAA  AAGTTCCCAA  TGCATGAGAT  GAGGGGCATT  TGCATTGCCT  CCCGTCACAC
301 TGCCCGCGAA  CCCCACCCC  ACCATAGCGT  CTGTGATAC  ATGGAGCGCG  AAGTCGAGAA
361 ACCTGTAATT  CCTGGTAACT  TTCAGGTACA  CAGTACGTAC  TGATCCTGGT  ATCAAACCTT
421 GCCTGCCGAG  TTTTCGACGG  AAAGAGGTGT  GAATTGTGAA  AGAGTCATAC  CAAATCACCC
481 GATTTTCATA  AAGCCCGAGT  CTTTTCTGTA  CATAAGCGAC  ACTCGAAGCG  GGCCTCATCT
541 TCATAGCCTG  ATAGCTTGTA  ATACTCCATC  CTCGTATCTC  ACTTGACCTT  GAGTTCAACC
601 CCACGTCAGA  CTTACCCCGA  CACATCGACG  GATTGGGGAA  CAGCACAATA  CCTGAAAAGC
661 GAGAAAACCA  AACAGAGGAA  AACACCATGG  AGACAACACT  TCCCCTCCCC  TTCCTCGTCG
721 GTGTCAGTGT  TCCTCCCGGA  CTGAATGACA  TCAAGGAGGG  CCTCAGCCGG  GAGGAAGTCT
781 CGTGTCTTGG  CTGCGTCTTC  TTCGAGGTCA  AGCCCAAGAC  CCTTGAGAAA  ATCGTGCGAT
841 TCCTCAAGCG  TCACAATGTC  GAATTTGAGC  CCTACTTCGA  TGTAAACAGCC  CTCGAGTCTA
901 TCGATGATAT  TATCACTCTT  CTGGACGCCG  GCGCCCGCAA  GGTGTTTGTG  AAGACCGAGC
961 AGTTGGCCGA  CCTCTCCGCA  TATGGCTCCC  GCGTTGCCCC  CATTGTCACT  GGAAGCAGCG
1021 CTGCTTTGCT  TTCCTCCGCC  ACCGAGAGCG  GCCTTTTGCT  CTCCGGCTTC  GATCAGACTG
1081 CCTCCGAGGC  TGCACAGTTT  CTGGAGGAGG  CCAGAGACAA  GAAAATTACC  CCCTTCTTCA
1141 TCAAGCCCGT  TCCTGGGGCC  GATCTCGAAG  AGTTTCATCA  GGTGCGCCGG  AAGGCTAACG
1201 CCATCCCCAT  CTTGCCATCC  ACTGGCTTGA  CAACAAAGAA  GGACGAGGCC  GGAAGGCTTG
1261 CCATCTCCAC  CATCTCTCG  AGCGTCTGGA  AGTCTGACCG  TCCCGATGGT  CTGCTCCCCA
1321 CCGTTGTCGT  TGATGAGCAC  GACACTGCTC  TGGGTCTGGT  CTACAGCAGT  GCCGAGAGTG
1381 TGAACGAGGC  CCTCAGGACA  CAGACTGGTG  TCTATCAGAG  CCGGAAGCGC  GGTCTCTGGT
1441 ACAAGGGTGC  TACTTCCGGA  GACACTCAGG  AGCTCGTCCG  CATCTCGCTT  GACTGCGATA
1501 ACGATGCTCT  CAAGTTTGTC  GTGAAGCAGA  AGGGTCGTTT  CTGCCACCTC  GATCAGTCCG
1561 GCTGCTTTGG  TCAGCTCAAA  GGCCTTCCCA  AGCTCGAGCA  GACTTTGATT  TCGAGGAAAC
1621 AGTCTGCCCC  CGAGGGCTCC  TACACTGCCC  GTCTCTTCTC  CGATGAGAAG  CTAGTCCGGG
1681 CCAAGATCAT  GGAGGAGGCT  GAGGAGCTCT  GCACCGCTCA  GACCCCCCAG  GAAATCGCCT
1741 TTGAGGCTGC  CGATCTCTTC  TACTTTGCTC  TTACCAGGGC  CGTTGCTGCC  GGC GTTACTC
1801 TTGCCGATAT  CGAAAGGAGC  CTTGACGCCA  AGAGCTGGAA  GGTCAGCGC  AGGACTGGAG
1861 ATGCTAAGGG  TAAGTGGGCT  GAGAAGGAGG  GCATCAAGCC  TGCGGCGTCC  GCTCCCCTG
1921 CCACTTCGGC  CCCTGTCACC  AAGGAGGCCG  CCCAGGAGAC  CACCCCTGAG  AAGATCACCA
1981 TGAGACGTTT  CGACGCCTCC  AAGGTCTCTA  CCGAGGAGCT  CGATGCTGCT  CTCAAGCGTC
2041 CTGCGCAAAA  GTCGTCCGAT  GCCATCTACA  AGATCATTTG  CCCCATCATC  GAGGACGTCC
2101 GCAAGAACGG  CGACAAGGCT  GTTCTGTGCT  ACACTCACAA  GTTCGAGAAG  GCTACCTCTC
2161 TTAGTAGCCC  CGTCCTGAAG  GCGCCCTTCC  CCAAGGAGCT  TATGCAGCTC  CCTGAGGAGA
2221 CCATTGCTGC  CATCGACGTG  TCCTTCGAGA  ACATCCGCAA  GTTCCACGCC  GCCCAGAAGG
2281 AGGAGAAGCC  CCTCCAGGTC  GAGACCATGC  CCGGTGTTGT  CTGCAGCCGT  TTCTCTCGTC
2341 CCATCGAGGC  CGTCGGCTGC  TACATCCCCG  GCGGTACCGC  CGTTCTCCCC  AGCACTGCCC
2401 TTATGCTGGG  TGTTCCCGCC  ATGGTCGCCG  GCTGCAACAA  GATTGTGTTT  GCCTCTCCTC
2461 CCCGCGCCGA  CGGAACCATC  ACTCCCGAGA  TTGTCCACGT  CGCTCACAA  GTTGGGGCCG
2521 AGTCCATCGT  GCTTGCCGGC  GGTGCCAGG  CCGTAGCTGC  CATGGCCTAC  GGCACCGAGA
2581 GCATACCAA  GGTCGACAAG  ATTCTCGGCC  CCGGTAACCA  GTTCGTCACT  GCTGCCAAGA
2641 TGTTCTGTCAG  CAACGACACC  AACGCTGCCG  TTGGGATTGA  CATGCCCGCT  GGCCCGTCCG
2701 AGGTGCTGGT  CATCGCTGAC  AAGGACGCCA  ACCCCGCGTT  CGTTGCCTCG  GATCTCCTGT
2761 CCCAGGCTGA  GCACGGCGTT  GACAGTCAGG  TCATCCTGAT  CGCTATTAA  CTCGACGAGG

```

# FIG. 7 continued

```

2821 AGCATCTTCA GGCTATTGAG GACGAGGTTT ACCGTCAGGC TATGGAGCTT CCTCGCGTCC
2881 AGATTGTCCG TGGCTCCATC GCCCACTCGA TCACCGTGCA GGTCAAGACC GTCGAGGAGG
2941 CCATGGAGCT CAGCAACAAG TACGCTCCTG AGCACTTGAT CCTCCAGATC AAGGAGGCCG
3001 AGAAAGCTGT CGATCTTGTC ATGAACGCTG GTAGTGTCTT CATTGGCGCT TGGACTCCTG
3061 AGTCCGTTGG CGATTACTCT GCTGGTGTTA ACCACTCGCT GCGTAAGTTA CATATCATAA
3121 ATAGCCCCGC TTCACAGATT CTTCTGCTAA CGTCAAGACA CATAGCTACC TATGGTTTTG
3181 GCAAGCAGTA CTCTGGCGTC AATCTCGCCT CGTTCGTCAA GCACATTACC AGCTCCAAC
3241 TGAAGCCGA GGGTCTCAA AACGTCGGCC AGGCTGTCAT GCAGTTGGCT AAGGTTGAGG
3301 AGCTCGAGGC TCACAGAAGG GCGGTCAGCA TCCGTCTTGA GCACATGAGC AAGAGCAACT
3361 AGACGGAAAT TCTTTTTTCGA AGTTGCAAAA AAAACAAGAA CAAAAGGATG TAGTGGGTTG
3421 ATGTATATCT GGGTCATTTT GGGCACATAG AGTAATGATA ACGAGTTTTG GACATTGTAC
3481 TGTTCTGTAC AGGCTGAAGA TCAGTACATG AATCTGTTGG TAAGTGTAGA GACCCAAACG
3541 TCCCTTGAGT TTTTCTCCCT GTTCCAGAGA GGTGCTCGTC CCTGGGTGTT TATTTTCATT
3601 ATTACATCAA CCTTTTATTT TATTTTATTT TTTATTTTAC TTTTTTTTCC TTTTTTTCAG
3661 ATCATGCGTA CATGAACGGG GGAAGCACAG ACGATCGAAA CGTGGATGTC ACAATGTGCG
3721 TGCAGTGATG CTGCATTGCA TGAAGCGCCC ATCTCAATAT ACTTGCAGTC TTGCGCGTTG
3781 CACGTGAAC TCCCAAACAA CCGAATAAAA GACGGCGAAA AATGAAGATA AAAAAAACC
3841 ATAATAAAAA TCGGAGGGAG TGTGGGAAAT GGTTTCTTTT AGCATTTAGA CCCCATAGCC
3901 GTGCACGCCC GGGTACAGAC AGGTTTCATCG ATGTTGACAT TGAAGGGAC ACCAGGTCTA
3961 TCTATTTTCT CTCTGTCTCT CTACCATAACA TCGGGACATC GGACATCTCG CTGTACCCCC
4021 CACACCCACA AAGTCTTATA AAAGCGCCAC ACCCGAGGAG GTTCGGTCCG CCCACGAAC
4081 TCCGTGCCTC CTGCCTGTT TACAGGGACC GAACGCTGGA GAAGCTTAGT TTCCTGACAT
4141 CCGGCCCTACC CGAGCAGGAA AAGGGACAGC TCATAGGCGA GGAGGGATTT GAAGATGGGG
4201 ACATTTTGGA TGATTCGAGA GGAGGAAC TA GGTACTGTAT CATGATAGTT CGGGGCGACA
4261 TCTTGGCTGG GACATTGTTA ATACCTCGAT ATGATGAAGT GGGAGGGAGT TTTTTCATGT
4321 CTGCCCCAAG TCCCACTAAT CTTTTTTTTT TTTTGTACCA ACACCAAGA TTCGGAGAAT
4381 AGTGAAGGA TTCGCATTCA CAAGTGGAAG TCTGAGGATC TTTTATATC TTTGTCTTCC
4441 GCGGACTGTT AACGATCCTA CAGCGAGCGA GCGAGCGGTC GGATGCGCTG ATCTGATAGG
4501 TGCAATATAC GGCCGCTTTC TCCGGTCGTG TAGTGTAAGC TCTGTCGGCA TAGTAGTACA
4561 CTAATAAAAC CCTTGCATTT CATGATCTGC TTGCTATTCA TTCCGAGTTA TTTAGTGGT
4621 CACATTTTCA GATTCACAGC CATCCATCCA TATGGAAAAA TCCATTCCCA TGCTTCCCTC
4681 CCCCCACTAT GTATGTGACC ACACGCTGCT GTCAGAATGC CAACGGTCTC AGGTACCCCTC
4741 GTCCGACTGT TTGGCATGGA GTTACATAACA CTACTAGTGT AGCCCCGGGC CAAGCTACCC
4801 CGTCAAATCT ATACATATCT ATAATGGGTT TCAGGTGTTT CGTTCGCTGT CAATCAAGTT
4861 TGAACATCA CTGGGGCCGT TGGACGGTGT ATTAGACCAT TGGCTCCCTC AGCTGGCGGC
4921 TGGGCGGTTG GGTCGGCAAT AACGGGACTG GACTTGAGAG GGACGAGGAG AGTCGGTTGG
4981 CTGCCTACAC TACACTACAA GCGTTCCCAC CTAACCGACG AGTCCCCTTT TCCATTTGTG
5041 TGCCTTAACC ATCATCTAGG GATGTCAGGG TTTGGCCGGA TCAGGGTATG TTTGGTTGAC
5101 TGTTGTCATG TCTGATTGGG TACATATCAT GGTAGGTGTC TCGAGAACAG TAGAGTACTC
5161 GGGCCTAGCG TTTGGATGAT TACGCGAGAT ATGAGTTGTA GGCCGCCATG CAGTTGCTTG
5221 CCCATAAGCA GAAGTTGCTT TGGGATATAT TTCTCGTCTT TCAAAGGTCA CGAGGTCCTG
5281 GGACGAGCGG CATCGCCATC CAAAGGGTTG AACATGAGAA ACCGGAATGG CCTTTGCGTT
5341 GAAATACAAA AAGTCAAGAA TAAAATCGCT TGAGGATAGG GACGTGGAAG CAAGCAAATA
5401 TGGTAAGGGA GGTACTGCTA TGTAAGTGCT CAGCAAACTG CCAATTTCTT GGCCCCAAG
5461 CAGCAGTTTG CTGTCAGTGC TGCTCGTGTC AGCCTTGGA GTGGAACCTA AACTGCTAAC
5521 ACAGCGCAAG TGCGCATGTA AAGATATTGT GGGAGGATCT GTATGGATGG ATGAGATTAC
5581 TGCTTGGTGT TGGTTGCGAG GCACTGCGGC TGTTAGGCTT TGCTGTGCCC CGTTCGACGA
5641 AGAAATACGC GGAACATAAA ATTGGATACC TAGACTTACT GCCTATGGGA GGTATCTACC
5701 GACGTAGCCG ACGGATTCTA GCAACATCCC GACTTTGCTT GTAGTGTACT ATGATAGCAG
5761 CACAGTGGGG TGTTGCTCCT TGTGAGCATG GGCTCTTTTT TTTTTTTTCC CCCTTCCCTA
5821 GGGCGTTGAC TGGACTTGCT CTATCGTTCC CAAGGTAGGT GCCCGTCATC GATTTTCCCA
5881 AGCCGTCTCC CGCCAGATTG TCGTCATAGT GTCATGATGA CCTCGGTGCG TGGGGCTGCG
5941 TGGTTACGGG GAGCTGGGAC CGCTAGGCCT CAGTGGTTGT GCCATTGAGC GTGGGTGTGT
6001 GGAGTAGCGG TAGAGGCGCT TGGAAAGTTG GCTAGCGGAA ACCCTGGAAT ATCTTGTACC

```

# FIG. 7 continued

6061	CTTCGATTCC	TTCTCGGGCT	GCCCATGTGC	TGAGGTGATG	CCGGGGATCT	GGCGCCAATC
6121	ATCCATTGAG	GTTCCCGCAG	CTTCCCGGTG	CCGCGCGCGG	GCGCAGTTGC	TCACAGGACA
6181	CACCTAGACG	CAGGGGCACA	GGGGCACCCT	TTGGTGTGCA	ACTGGGTACC	TGGTAGCTGT
6241	AGCAAGCACT	CCACCGTCTG	TGCAATCCCC	CAATCCACGG	CAGGAACCTA	GCACCGCCGC
6301	GGCACCAGT	GAGCGAATCC	ATCCGCATTG	GATCCCAATT	CTTGCCCTTG	CCATCCTTCT
6361	TTCTTCCCAC	TTGGCGCAAC	CAACACTTCC	CTTGGTCTGG	GTACTCGTGT	TGATCTTCAC
6421	TCTCTTTTTT	TCTTGGGCGA	CCGACTTTTT	ATATCCGTCC	TTGCTTCCCC	CTGGCCGTTG
6481	TCGTTCTTTC	TACAACCTACC	TTCCGTTTAT	TATCCCCCTT	CTTGGTTCGG	TCGAGGACCC
6541	AAAAACAGAA	CAATTCCGGC	TCTTCCAGGT	GGCTTGGGTG	CGACTGTTTA	GCTCTTGACC
6601	ACTAGCCGCT	TACCTTCTCT	TGATGTTTAT	ATTTGGATAT	CATTGAACTA	CTCTTTCTTG
6661	AAACGGCAGA	CGAACGGAAC	AGTCCCTACG	GTTTATTAGC	GATATACGTT	GTACTGATAT
6721	CCTGAGCAAG	AAGAGGCAAA	TTATCAATTA	TGCATCTCCC	ATCGTCGCTG	CTCATCGCAG
6781	CTCCCTTGCT	CGCCAATGTA	TCGGCCGAAC	CGATTAGGAT	ACCCCAACGC	GATGTTCTCC
6841	GTGGTATCAA	CATCACAGCA	ACTTGCCGTT	CGAGCACTAC	CGAATTGCGC	CAGCGGTGGA
6901	TATGCCCTTG	CCGTTGTAGA	CTGTCCCAAG	ACCAAGCCGA	CGCTCCGGAA	GGCCGTGGAT
6961	TTGTCGAACG	AGGAGAAGAA	CTGGTTGTCT	ATCCGGAGGA	AGAACACCAT	CCAGCCCATG
7021	AGGGACCTAC	TGAAGAGGGC	CAACATCACT	GGGTTTCGATT	CCGAAACTTT	CATGAATGAG
7081	GCCGCCAACA	ACGTCCTCGCA	ACTGCCCAAT	GTCGCCATTG	CCATTTTCAGG	AGGCGGCTAT
7141	CGTGCCCTCA	TGAACGGCGC	CGGCTTCGTT	GCTGCTGCGG	ATAACCGGAT	TCAAATACC
7201	ACGGGCGCAG	GTGGTATTGG	AGGCTTGTTG	CAGTCCAGCA	CATATTTGTA	TGTAAACCA
7261	TGCCCTTCTT	TGGTTCTTCT	TATCTCGTTT	TCGAGTGTCA	ACTGCGCCAG	TTTCGACGTTG
7321	GGCGGCTGTG	GACGACCTTG	CTGGTGAACA	TGTCTTGGAC	TCCATGCCCC	TTTTTTCGTT
7381	CCCTAAATC	CCAAAAA	AAAAA	AAAAA	AAAAA	AAAAATTCGAG
7441	GACCGTGACT	GTAATTGCT	AACGCAACTC	TAGGGCCGGA	CTTTCTGGTG	GTGGCTGGCT
7501	TGTCGGCAGT	TTGTTCTCCA	ACAACCTCAG	TAGCATTGAG	ACCCTGCTGA	GCGAGAACAA
7561	AGTCTGGGAC	TTTGAGAACT	CCATCTTTAA	AGGACCCAAG	GAGGCTGGCC	TTAGTACTGT
7621	CAACCGTATC	CAGTACTGGT	CCGAAGTGGC	AAAGGAAGTT	GCGAAGAAGA	AGGATGCTGG
7681	CTTCGAGACA	AGTATAACAG	ACTACTGGGG	CCGAGCATTG	AGTTACCAAC	TGATCGGAGC
7741	CGATATGGGC	GGCCCGGCTT	ACACCTTCTC	CAGCATTGCC	CAGACCGACA	ACTTCCAGAA
7801	GGCCGAAACG	CCGTTCCCTA	TTCTGGTAGC	TGACGGCCGC	GCGCCTGGAG	ACACCATCAT
7861	CTCCCTCAAT	GCTACCAACT	ACGAGTTCAA	CCCGTTTCGAG	ACGGGTAGCT	GGGACCCGAC
7921	CGTCTATGGC	TTTGCGCCGA	CCAAGTACCT	CGGCGCCAAC	TTTCAGCAACG	GCGTGATCCC
7981	ATCGGGAGGC	AAGTGCGTTG	AGGGTCTCGA	CCAAGCCGGC	TTTCGTATGG	GCACCAGCAG
8041	CACGCTCTTC	AACCAGTTCC	TTTTGGCCAA	CATCTCCAGC	TACGACGGTG	TTGCCAGACG
8101	TGCTCATCGA	GGCCGTGACT	TCTGTCTCTA	AGGAAATCGG	CGCCAAGAGG	ACGACGTCTC
8161	CCAAATCATC	CCTAATCCGT	TCCTGGACTG	GAACAACCGG	ACCAACCCCA	ACGCCGACAC
8221	GCTCGAGCTC	GACCTGGTCG	ACGGCGGCGA	AGATCTGCAG	AATATTCCGC	TCAACCCGCT
8281	CACCCAACCC	GTGCGCGCCG	TCGACGTCAT	CTTCGCTGTC	GACTCGTCCG	CCGACGTGAC
8341	AAACTGGCCC	AATGGCACCG	CCCTGCGCGC	CACCTACGAG	CGCACTTTCG	GCTCTATTTC
8401	CAACGGGACA	CTCTTCCCCT	CGATCCCCGA	CGACTGGACG	TTTATAAACC	TAGGCCCTCAA
8461	CAACCGCCCC	TCTTTCTTCG	GCTGCGATGT	TAAGAACTTT	ACCTTGAACG	CCAACCAAAA
8521	GGTTCCTCCC	TTAATCGTCT	ATGTCCCCAA	CGCGCCCTAT	ACCGCGCTGA	GCAACGTGTC
8581	CACCTTCGAT	CCGTCATACA	CGATGTCTCA	GCGCAACGAC	ATCATCGGCA	ACGGATGGAA
8641	CTCAGCCACG	CAGGGAAACG	GCACGCTGGA	TTCGGAGTGG	CCCACTTGCG	TCGCCTGCGC
8701	GGTTATCAGC	AGGAGCTTAG	ATCGGTTGGG	CAGGCAGACG	CCAGCCGCGT	GCAAGACTTG
8761	CTTTGACAGG	TATTGCTGGA	ATGGCACAGT	GAACCTCAAA	GATACGGGGG	TTTACATGCC
8821	TGAGTTCAAG	ATTGCGGATG	CGCATGCCCT	GGACTCGGGT	GCTGTTGCTA	TCGGAAAGAT
8881	GGTGAATGTC	TGGTCGTCGG	TTGTGGTGGG	AGTTGTGGCG	GCTACTTTGT	TGTTGTAGGG
8941	GTAGGGGAGA	CGTGATGATA	TTCCAGTCTG	ATGAAGTTGA	GACTGGACTG	GAGATCGCCA
9001	AGGATGCGGA	GGGAAAGGAA	TGCGTGGTGT	TAATGTCATG	ATGGATGAAG	AGTCATGGAT
9061	CATGGAACGA	CGGGGCGGGG	ATATTGGATG	ATGGATATAC	CACACTGCAT	GCATGCTCTA
9121	TTGATAGTAT	GCTTTGGCAT	TTACGTTTAA	CAATCAATTG	CTCCATCCTG	ATGTTCTATC
9181	TTTTTCGACA	ATGGATTGAT	ACTACTCCTG	TTGCTTCGCT	CTTGAGGTTG	GAAGGACTTG
9241	AGGTTGGAAG	GACTTGAGGT	TGTTTGTTCT	GAGGGAGGTT	ATCGAAGTAT	CATCTGTGCT

## FIG. 7 continued

```
9301 GATGCCGATT GATAGACTGT CCTCTTCTTC GAGGCAACGA ACGGTCGGAT GAGCCTCTTT
9361 AATCATGATG CTCAGTGCCA CAAAAAGGCT CCAGCACAGC TGCCCACACC TTTCTTGCCT
9421 CGCCGTTCCT TCCTTTTTCT TTTCCCCTGT TTCCTTTCTT CCTTTCCATC TCATCCCGTA
9481 CCAGAGTGCC CACCGGGTAT ATATATTACC TCCTTGGCCG TTCTCCTTTG ACCAATAAAT
9541 CGCTTGGTCG AGTGGCGTAA CGGTTTACCG TCTACACTTA TCACTCAAAC CAAACCAAAC
9601 CATCGAAGAA GTGACCTATC GGTTTCGAGG AACGGTGATG TTCTTACGAC CAAGTTAACC
9661 CAAAGAGCGT TCCACATCGT TGAACCGTCT CCTCCAGTTG GATCTGTTTA ACTTCCGCAG
9721 CGACTGAAGA AGGTATCACT TTTTTTTTGG TTCCAAAAAA AAAAAAAAAA ATTAC
```

# FIG. 8

Nucleotide sequence of the *his-3 cog<sup>E</sup> lpl* region of linkage group I in the StLawrence wild type strain of *Neurospora crassa*. This differs from that in the Lindegren strain in many positions, summarised in figure 5. The coordinates of relevant features are given in the text. This sequence contains the weak recombinator *cog<sup>E</sup>* and also the remnant of a transposable element *Guest* within the replaceable sequence 3' of *his-3*. StLawrence strains carry *rec-2<sup>+</sup>* which prevents the initiation of recombination at *cog*.

```

1  ACCGGGAATC GTAGCGGGCG CTAAGGCCAA GCCGCGGCAC GGGTCACTGA CCCAATGCAG
61  CGCATTCCGT CAGCAACTGA AGTGGATGTA CAAGTACATA GTAGTAGATC GCAACTGGAG
121 ATCACTCGCA CCGTGCCGCA GAACAAGGGC GACGAGCCTC AGGGCAGTTT AGCCTGCCGT
181 AACAGCACAG ACCATAGCTT ATTTTCACCT GGGCGGGCGG GCGACGGCGG CACTGACATC
241 GGCAAGGCGG CATCAAGCAA CCCCTCTGTT GCTTGCCAGC TGCCGGCCAA CGTCAGCGGT
301 ACAAGGAGAA ATCTGGAAGG AAAGACTTCT GGCACCGACA GGATGGCACG CGGGAAAAGT
361 TCCCAATGCA TGAGATGAGG GGCATTGCA TTGCCTCCCG TCACCCAGTG CGAACCCCAA
421 CCCCAACATA GCGTCTGTCT ATACATGGAG CGCGAAGTCG AGAAACCTGT AATTCCCTGGT
481 AACTTTCAGG TACACAGTAC GTACTGATCC TGGTATCAAA CCTTGCCCTGC CGAGTTTTTCG
541 ACGGAAAGAG GTGTGAATTG TGAAAGAGTC ATACCAAATC ACCCGATTTT CATAAAGCCC
601 GAGTCTTTTC TGTACATAAG CGACACTCGA AGCGGGCCTC ATCTTCATAG CCTGATAGCT
661 TGTAATACTC CATCCTCGTA TCTCACTTGA CCTTGAGTTC AACCCACAGT CAAACTTCAC
721 CCGACACATC GACGGATTGG GGAACAGCAC AATACCTGAA AAGCGAGAAA ACCAAACAGA
781 GGAAAACACC ATGGAGACAA CACTTCCCTT CCCCTTCCCTC GTCGGTGTCA GTGTTCTCTC
841 CGGACTGAAT GACATCAAGG AGGGCCTCAG CCGGGAGGAA GTCTCGTGTG TTGGCTGCGT
901 CTTCTTCGAG GTCAAGCCCC AGACCCTTGA GAAAATCCTG CGATTCCCTCA AGCGTCACAA
961 TGTCGAATTT GAGCCCTACT TCGATGTAAC AGCCCTCGAG TCTATCGATG ATATTATCAC
1021 TCTTCTGGAG GCCGCGCCCC GCAAGGTGTT TGTCAAGACC GAGCAGTTGG CCGACCTCTC
1081 CGCATATGGC TCCC CGTGGT CCCCATTGT CACTGGAAGC AGCGCTGCTT TGCTTTCCTC
1141 CGCCACCGAG AGCGGCCTTT TGCTCTCCGG CTTCGATCAG ACTGCCTCCG AGGCTGCACA
1201 GTTTCTGGAG GAGGCCAGAG ACAAGAAAAT TACCCCTTC TTCATCAAGC CCGTTCCTGG
1261 GGCCGATCTC GAACAGTTCA TCCAGGTCGC CGCCAAGGCT AACGCCATCC CCATCCTGCC
1321 ATCCACTGGC TTGACAACAA AGAAGGACGA GGCCGGCAAG CTTGCCATCT CCACCATCCT
1381 CTCGAGCGTC TGGAAGTCTG ACCGTCCCGA TGGTCTTCTC CCCACCGTTG TCGTTGATGA
1441 GCACGACACT GCTCTGGGTC TGGTCTACAG CAGTGCCGAG AGTGTGAACG AGGCCCTCAG
1501 GACACAGACT GGTGTCTATC AGAGCCGGAA GCGCGGTCTC TGGTACAAGG GTGCTACTTC
1561 CGGAGACACT CAGGAGCTCG TCCGCATCTC GCTTGACTGC GATAACGATG CTCTCAAGTT
1621 TGTCGTGAAG CAGAAGGGTC GTTCTGCCA CCTCGATCAG TCCGGCTGCT TTGGTCAGCT
1681 CAAAGGCCTT CCCAAGCTCG AGCAGACTTT GATTTCGAGG AAACAGTCTG CCCCCGAGGG
1741 CTCCTACACT GCCCGTCTCT TCTCCGATGA GAAGCTAGTC CGGGCCAAGA TCATGGAGGA
1801 GGCTGAGGAG CTCTGCACCG CTCAGACCCC CCAGGAAATC GCCTTTGAGG CTGCCGATCT
1861 CTTCTACTTT GCTCTTACCA GGGCCGTTGC TGCCGGCGTT ACTCTTGCCG ATATCGAAAG
1921 GAGCCTTGAC GCCAAGAGCT GGAAGGTCAA GCGCAGGACT GGAGATGCTA AGGGTAAGTG
1981 GGCTGAGAAG GAGGGCATCA AGCCTGCGGC GTCCGCTCTC GCTGCCACTT CGGCCCCTGT
2041 CACCAAGGAG GCCGCCAGG AGACCACCCC TGAGAAGATC ACCATGAGAC GTTTCGACGC
2101 CTCCAAGGTC TCTACCGAGG AGCTCGATGC TGCTCTCAAG CGTCTGCGC AAAAGTCGTC
2161 CGATGCCATC TACAAGATCA TTGTCCCAT CATCGAGGAC GTCCGCAAGA ACGGCGACAA
2221 GGCTGTTCTG TCGTACACTC ACAAGTTCGA GAAGGCTACC TCTCTTACTA GCCCCGTCTT
2281 GAAGGCGCCC TTCCCCAAGG AGCTTATGCA GCTCCCTGAG GAGACCATTG CTGCCATCGA
2341 CGTGTCTTTC GAGAACATCC GCAAGTTCCA CGCCGCCAG AAGGAGGAGA AGCCCCTCCA
2401 GGTCGAGACC ATGCCCGGTG TTGTCTGCAG CCGTTTCTCT CGTCCCATCG AGGCCGTCGG
2461 CTGCTACATC CCCGGCGGTA CCGCCGTTCT CCCCAGCACT GCCCTTATGC TGGGTGTTCC
2521 CGCCATGGTC GCCGGCTGCA ACAAGATTGT GTTCGCCTCT CCTCCCCGCG CCGACGGAAC
2581 CATCACTCCC GAGATTGTCC ACGTCGCTCA CAAGGTTGGG GCCGAGTCCA TCGTGCTTGC
2641 CGGCGGTGCC CAGGCCGTAG CTGCCATGGC CTACGGCACC GAGAGCATCA CCAAGGTCTA
2701 CAAGATTCTC GGCCCCGTA ACCAGTTCGT CACTGCTGCC AAGATGTTCTG TCAGCAACGA

```

# FIG. 8 continued

```

2761 CACCAACGCT GCCGTTGGTA TTGACATGCC CGCTGGCCCCG TCCGAGGTGC TGGTCATCGC
2821 TGACAAGGAC GCCAACCCCG CGTTCGTTGC CTCGGATCTC CTGTCCCAGG CTGAGCACGG
2881 CGTTGACAGT CAGGTCATCC TGATCGCTAT TGACCTCGAC GAGGAGCATC TTCAGGCTAT
2941 TGAGGACGAG GTTCACCGTC AGGCTACGGA GCTTCCTCGC GTCCAGATTG TCCGTGGCTC
3001 CATCGCCAC TCGATCACCG TGCAGGTCAA GACCGTCGAG GAGGCCATGG AGCTCAGCAA
3061 CAAGTACGCT CCTGAGCACT TGATCCTCCA GATCAAGGAG GCCGAGAAGG CTGTCGATCT
3121 TGTCATGAAC GCCGGTAGTG TCTTCATTGG CGCCTGGACT CCTGAGTCCG TTGGCGATTA
3181 CTCTGCTGGT GTTAACCACT CGCTGCGTAA GTTACATATC ATAAATAGCC CCGCTTCACA
3241 GATTCTTCTG CTAACGTCAA GACACATAGC TACCTATGGC TTTGGCAAGC AGTACTCTGG
3301 CGTCAATTTT GCCTCGTTTCG TCAAGCACAT TACCAGCTCC AACTTGACTG CCGAGGGTCT
3361 CAAAAACGTC GGCCAGGCTG TCATGCAGTT GGCTAAGGTT GAGGAGCTCG AGGCTCACAG
3421 AAGGGCGGTC AGCATCCGTC TTGAGCACAT GAGCAAGAGC AACTAAACGG AAATTCTTTT
3481 CGAAGTAGCA AAAAAAAAAA AAAAAACAA GAACAAAAGG ATGTAGTGGG TTGATGTATA
3541 TCTGGGTCAT TTTGGGCACA TAGAGTAATG ATAACGAGTT TTGGACATTG TACTGTTCTG
3601 TACAGGCTGA AGATCAGTAC ATGAATCTGT TGGTAAGTGT GGAGACCCAA ACGTCCCTTG
3661 AGTTTTTCTC CCTATTCCAG AGGTGCTCGT CCCTGGGTGT TTATTTTCAT TATTACATCA
3721 ACCTTTTTTT TTTTTTTTTT TTTTTCAGAT CATGCGTACA TGAACGGGGG AAGCACAGAC
3781 GATCGAAACG TGGATGTCAC AATGTCGCTG CAGTGATGCT GCATTGCATG AAGCGCCCAT
3841 CTCAATATAC TTGCAGTCTT GCACGTTGCA TGTGAAC TTC AACAACCC GAATAAAAGA
3901 CGGCGAAAAA TGAAGATAAA AAAAAACCAT AAAAAAATC AGAGGGAGTG TGGGAAATGG
3961 TGTCTTTTAG CATTCAGACC CCATAGCCGT GCACGCCCGG GTACAGACAG GTTCATCGAT
4021 GTTGACATTG ACTGGGACAC CAGGTCTATC TATTTTATCT CCTGTCTCT ACCATACATC
4081 GGGACATCGG ACATCTTGCT GTACCCCCCA CACCCACAAA GCCTTATAAA AGCGCCACAC
4141 CCGAGGAGGT TCGGTCGGCC CCACGAACTC TGTGCCCTCC TGCTGTTTA CAGGGACCGA
4201 ACGCTGGAGA ATCTTACTAG TTTCTGACA TCCGGCCTAC CCGAGCAGGA AAAGGGACAG
4261 CTCATAGGCG AGGAGGGATT TGAAGATGG AACATTTTGG GTGATTGAG AGGAGGAAC
4321 AGGTACTGCA TCATGATAGT TCGGGGCAGC ATCTTGGCTG GGACATTGTT AATACCTCGA
4381 TATGATGAAG TAGGAGGGAG TTTTTCGCTG TCTTGCCGAA GTCCAGAGAT CTGTTTTATT
4441 TTATTTTTTA TGGATGTAGT GTATCAACAC CCAAGATTCG GAGAATAGTA CTAGGATTCG
4501 CATTTACAAG TGGAAGTCTT GAGAATCGTT GTATATCCTT GTCTTCCTCG GAATGTTAAC
4561 AATCTACAG CGAGCGAGCG AGCGGTGCGA TGCGCTGATC TGATAGGCGC AATATACGGC
4621 CGCTTTCTCC GGTGCTGTAG TGTAAGCTCT GTGGGCATAG TACACTAAAA AAACCCTTGC
4681 ATTTTCATGAT CTGCCTGCTA TTCATTCGGA GCTATTTTCA TGGTCACATT TCGAGGAAGA
4741 AAGAAAGCAA CTAAGATTCA CAGCCATCCA TCCATCCATA TGGAAGAATA ATCCATTCCC
4801 ATGTTCCCTC CCCCCACTA TGATGTGAC CACACGCTGC TGTCAGAATG CCAACGGTCT
4861 CAGGTACCTT CGTCCGACTG TTTGGCATGG AGTTACATAC ACTACTAGTG TAGCCCCGGG
4921 CCAAGCTACC CCGTCAAATC TATACATATC TATAACGGGT TTCAGGGGTT TCGTTCGCTG
4981 TCAATCAAGT TTGAAACATC ACTGGGGCCG TTGGACGGTG TATTAGACCA TTGGCTCCCT
5041 CAGCTGTTTG GCGGCTGGGC GGCTGGGTCA AACGGCAATA ACGGGACTCG AGAGGGACGA
5101 GGAGAGTCGG TTGGCTGGCT GCAATACAAG CGTTCCCAAC TAACCAACGA GTCCCGTTTT
5161 CCATTTGTGT GCCTAACCAT CATCTAGGGA TGTGAGGGT TGCCCGATC AGGGTATGTT
5221 TGGTTGACTG TTGTCATGTC TGATTGGGTA CATATTATGG TAGGTGTCTC GAGAACAGTA
5281 GAGTACTCGG GCCTAGCGTT TGGATGATTA CGCGAGATAT GAGTTGTGGG CCGCCATGCA
5341 GTTGCTTGTG CATAAGCAGA AGTTGCTTTG GGATATATTT CTCGTCTTTC AAAGTTCACG
5401 AGGTCTTGGG ACGAACGGCA TCGCCATCCA AAGGGTTGAA CATGAGAAAC CTGAATGGCC
5461 TTTGCGTTGA AATACAAAAA GTCAAGAACA AAATCGCTTG AGGATAGGGA CGTGGAAGCA
5521 AGCAAATATG GTAAGAGAGG TATACATCAA CCCTGGTTCA ATTGTTAGCG TGGTCTTCC
5581 TCCACGTCCT CGTTCATGAC GGTTAACAGT ACCAGGCTAA CAATTAAACC AGGGTTGATG
5641 TGTAAGTATA TGTAAGTGCT CAGCAAACTG CCAATTTCTT TGGCCCCAAG CAGCAGTTTG
5701 CTGTCAGTGC TGCTCGTGTC AGCCTTGGTA GTGGAACCTA AACTGCTAAC ACAGCGCAAG
5761 TGCGCATGTA AAGATATTGT GGGAGGATCT GTATGGATGG ATGAGATTAC TGCTTGGTGT
5821 TGGTTGCGAG GCACTGCGGC TGTTAGGCTT TGCTGTGCCC CGTTCGACGA AGAAATACGC
5881 GGAACATAAA ATTGGATACC TAGACTTACT GCCTATGGGA GGTATCTACC GACGTAGCCG
5941 ACGGATTCTA GCAACATCCC GACTTTGCTT GTAGTGTACT ATGATAGCAG CACAGTGTG

```

# FIG. 8 continued

```

6001 CTCCTTGTGA GAATGGGCTC TTTTTTTTTT TCCCCCTTCC CTAGGGCGTT GACTGGACTT
6061 GCTCTATTGT TCCCAAGGTA GGTGCCCCGTC ATCGATTTTC CCAAGTCTCC CGCCAGATTG
6121 TCGTCATAGT GTCATGATGA CCTCGGTCGC TGGGGCTGCG TGGTTACGGG GAGCTGGGAC
6181 CGCTAGGCCCT CAGTGGTTGT GCCATTGAGC GTGGGTGTGT GGAGTAGCGG TAGAGGCGCT
6241 TGGAAGTTGT GCTAGCGGAA ACCCTGGAAT ATCTTCTACC CTCGATTCCCT TCTCGGGCTG
6301 CCCATGTGCT GAGGTGATGC CGGGGATCTG GCGCCAATCA TCCATTGAGG TTCCCGCAGC
6361 TTCCCGGTGC CGCGCGCGGG CGCAGTTGCT CACAGGACAC ACCTAGACGC AGGGGACAG
6421 GGGCACCATT TGGTGTGCAA CTGGGTACCT AGCTGTAGCA AGCACTCCAC CGTCTGTGCA
6481 ATCCCCCAAT CCACGGCAGG AACTTCGCAC CGCCGCGGCA CCGAGTGAGC GAATCCATCC
6541 GCATTGGATC CCAATTCTTG CCCTTGCCAT CCTTCTTTCT TCCCCTTGG CGCAACCAAC
6601 ACTTCCCTTG GTCTGGGTAC TCGTGTGAT CTTCACCTCT TTTTTTCTT GGGCGACCGA
6661 CTTTTTATAT CCGTCCCTTG TCCCCCTGG CCGTTGTCGT TCTTTCTACA ACTACCTTCC
6721 GTTCATTATC CCCTTTCTTG GTTCGGTCGA GGACCCAAAA ACAGAACAAT TCCGGCTCTT
6781 CCAGGTGGCT TGGGTGCGAC TGTTTAGCTC TTGACCACTA GCCGCTTACC TTCTCTTGAT
6841 GTTTTTATTT GGATATCATT AAATACTCT TTCTTGAAAC GGCAGACGAA CGGAACAGTT
6901 CCTACGGTAT ATTAGCGATA TACGTTGTAC TGATATTCTG AGCAAGAAGA GGCAAATTAT
6961 CAATTATGCA TCTCCCTTCG TCGCTGCTCA TCGCAGCTCC CTTGCTCGCC AATGTATCGG
7021 CCGAACCCAT TAGGATACCC CAACGCGATG TTCTCCGTGG TATCAACATC ACAGCAACTT
7081 GCCGTTTCGAG CACTACCGGA TTCGCCCAGC GGTGGATATG CCCCTGCCGT TGTAGACTGT
7141 CCCAAGACCA AGCCGACGCT CCGGAAGGCC GTGGATTTGT CGAACGAGGA GAAGAACTGG
7201 TTGTCGATCC GGAGGAAGAA CACCATCCAG CCCATGAGGG ACCTCCTGAA GAGGGCCAAC
7261 ATCACTGGGT TCGATTCCGA GACATTTATG AATGAGGCCG CCAACAACAT CTCGCAACTG
7321 CCCAATGTCG CCATTGCCAT TTCAGGAGGC GGCTATCGTG CCCTCATGAA CGGCGCCGGC
7381 TTCGTTGCTG CTGCGGATAA CCGAATTCAA AATACCACGG GCGCAGGTGG TATTGGAGGC
7441 TTGTTGCAGT CCAGCACATA TTTGTATGTA AAGTGGTTCT TCTTATCTCG TTTTCGATG
7501 TCAACTGCGC CAGTTCAGAG TTGGGCGGCT TTGCGGACCG TTGCTGGTGA ACATGCTTTG
7561 GACTCCATGC CCCTTCTTCG TTTCTCAA AA TCAAGAAGTC GAGGACCGTG ACCGTAAATC
7621 GCTAACGCAA CTCTAGGGCC GGACTTTCTG GTGGTGGCTG GCTTGTGCGC AGTTTGTCTT
7681 CCAACAACCTT CAGCAGCATT GAGACCCTGC TGAGCGAGAA CAAAGTCTGG GACTTTGAGA
7741 ACTCCATCTT TAAAGGGCCC AAGGAGGCTG GCCTTAGTAC TGTCAACCGC ATTCAGTACT
7801 GGTCCGAAGT GGCAAAGGAA GTTGCCAAGA AGAAGGATGC TGGCTTCGAG ACAAGTATAA
7861 CAGACTACTG GGGCCGAGCA TTGAGTTACC AACTGATCGG AGCCGATATG GGGCGCCCGG
7921 CTTACACCTT CTCCAGCATT GCCCAGACCG ACAACTTCCA GAAGGCCGAA ACGCCGTTCC
7981 CTATTCTGGT AGCTGACGGC CGCGCGCCTG GAGACACCAT CATCTCCCTC AATGCTACCA
8041 ACTACGAGTT CAACCCGTT CAGACGGGTA GCTGGGACCC GACCGTCTAT GGCTTTGCGC
8101 CGACCAAGTA CCTCGGCGCC AACTTCAGCA ACGGCGTGAT CCCATCGGGA GGCAAGTGCG
8161 TTGAGGGTCT CGACCAAGCC GGCTTCGTCA TGGGCACCAG CAGCACGCTC TTCAACCAGT
8221 TCCTTTTGGC CAACATCTCC AGCTACGACG GTGTTGCCCC ACGTGCTCAT CGAAGCCGTG
8281 ACTTCTGTCC TCAAGGAAAT CGGCGCCAAG AGGACGACGT CTCCCAAATC ATCCCTAATC
8341 CGTTCTTGGA CTGGAACAAC CGGACCAACC CCAACGCCGA CACGCTCGAG CTCGACCTGG
8401 TCGACGGCGG CGAAGATCTG CAGAATATTC CGCTCAACCC GCTCACCCTA CCCGTGCGCG
8461 CCGTGACGT CATCTTCGCT GTCGACTCGT CCGCCGACGT GACAACTGG CCCAATGGCA
8521 CCGCCCTGCG AGCCACCTAC GAGCGCACTT TCGGCTCTAT TTCCAACGGG ACACTCTTCC
8581 CCTCGATCCC CGACGACTGG ACGTTTATAA ACCTAGGCCT CAACAACCGC CCCTCTTTCT
8641 TCGGCTGCGA TGTTAAGAAC TTTACCTTGA ACGCCAACCA AAAGGTTCCC CCCTTAATCG
8701 TCTATGTCCC CAACGCGCCC TATACCGCGC TGAGCAACGT GTCCACCTTC GATCCGTCAT
8761 ACACCATGTC TCAGCGCAAC GACATCATCG GCAACGGATG GAACTCAGCC ACGCAGGGAA
8821 ACGGCACGCT GGATTGCGAG TGGCCCACTT GCGTCGCTG CGCGGTTATC AGCAGGAGCT
8881 TAGATCGGTT GGGCAGGCAG ACGCCAGCCG CGTGCAAGAC TTGCTTTGAG AGGTATTGCT
8941 GGAATGGCAC AGTGAACCTA AAAGATACAG GGGTTTACAT GCCTGAGTTC AAGATTGCGG
9001 ATGCGCATGC CCTGGACTCG GGTGCTGTTG CTATCGGAAA GATGGTGAAT GTCTGGTCGT
9061 CGGTTGTGGT GGGAGTTGTG GCGGCTACTT TGTTGTTGTA GGGGTAGGGG AGACGTGATG
9121 ATATTCCAGT CTGATGAAGT TGAGACTGGA CTGGAGATCG CCAAGGATGC GGAGGGAAAG
9181 GAATGCGTGG TGTTAATGTC ATGATGGATG AAGGGTCATG GATCATGGAA CGACGGGGCG

```

## FIG. 8 continued

```
9241 GGGATATTGG ATGATGGATA TACCACACTG CATGCATGCT CTATTGATAA TATGCTTTGG
9301 CATTTACGTT TAACAATCAA TTGCTCCATC CTGATGTTCT ATCTTTCGAC ACTGGATTGA
9361 TACTACTCCT GTTGCTTCCC TCTTGAAGTT GGAAGGACTT GAGGTTGGAA GGACTTGAGG
9421 TTGTTTGTTT TGAGGGAGGT TATCGAAGTA TCATCTGTGC TGATGCCGAT CGATAGACTG
9481 CCCTCTTCTT CGAGGCAACG AACGGTCGGA TGAGCCTCTA ATCATGATGC TCAGTGCCAC
9541 AAAAAGGCTC CAGCACAGCT GCCCACACCT TTTTGCCTC GTCGCTCCTT CCTTTTTTTC
9601 CCCCCCTTTC TTCCTTTCCA TCTCATCCCG TACCAGAGTG CCCACCGGGT ATATATATTA
9661 CCTCCTTGGC CGTTCTCCTT TGACCAATAA ATCGCTTGGT CGAGTGGCGT AACCGTTTAC
9721 CGTCTACACT TATCACTCAA ACCAAACCAA ACCATCGAAG AAGTTACCTA TCGGTTTCGAG
9781 GGAACGGTGA TGTTCTTACG TTCAAGTTAA CCCAAAGAGC GTTCCACATC GTTGAACCGT
9841 CTCCTCCAGT TCTTGGATCT GTTTAACTTC CGCAGCGACT GAAGAAGTAA TCACTTTTTT
9901 TTTTTTTGGT TCCAAAAAAA AAAAAAAAAA TTAC
```

## FIG. 9

Construction of the components of the sequence diversification cross: Parent (variant 1) and Parent (variant 2). For convenience, plasmid sequences are shown as linear. The cross hatched region in the chromosome is dispensable. Stippled sequence in the plasmid indicate the multiple cloning site for inserting foreign DNA. Crossovers in region 1 and region 2 insert the foreign sequence to be diversified into chromosome 1 of *Neurospora crassa* adjacent to the recombination hotspot *cog*. Parent (variant 2) containing a version of the foreign sequence with multiple differences from that in parent (variant 1) is similarly constructed. Parent (variant 1) and parent (variant 2) are crossed and conversion events (stippled arrow) initiated (X) at *cog*<sup>L</sup> recombine the sequence differences in variant 1 and variant 2 to form new combinations. Sequences are identical except for those that distinguish variant 1 and variant 2. *rec-2* on linkage group V permits *cog*<sup>L</sup> to be active. For simplicity, genes not directly related to the diversification are omitted. See text for further details.

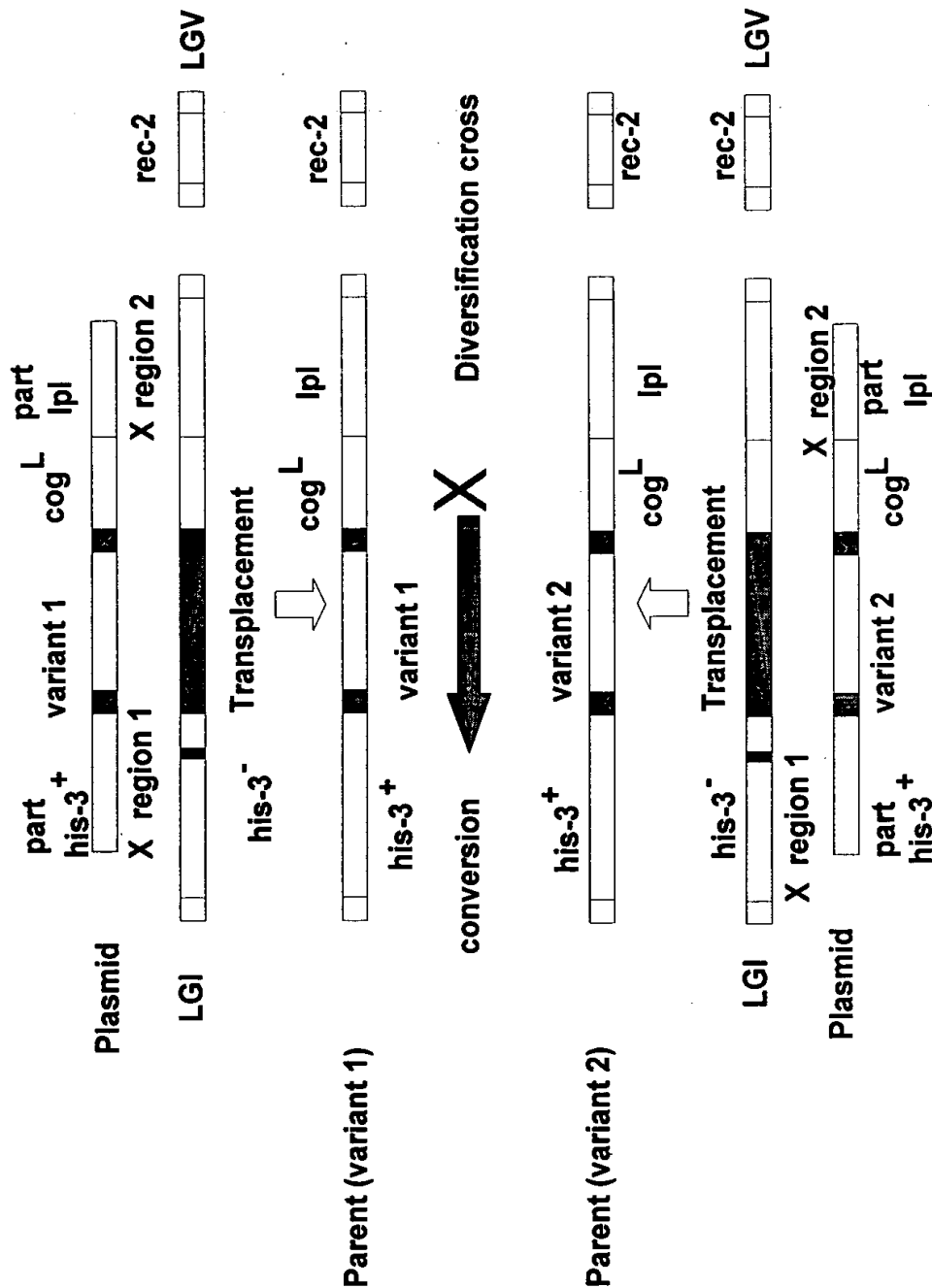




FIG. 11

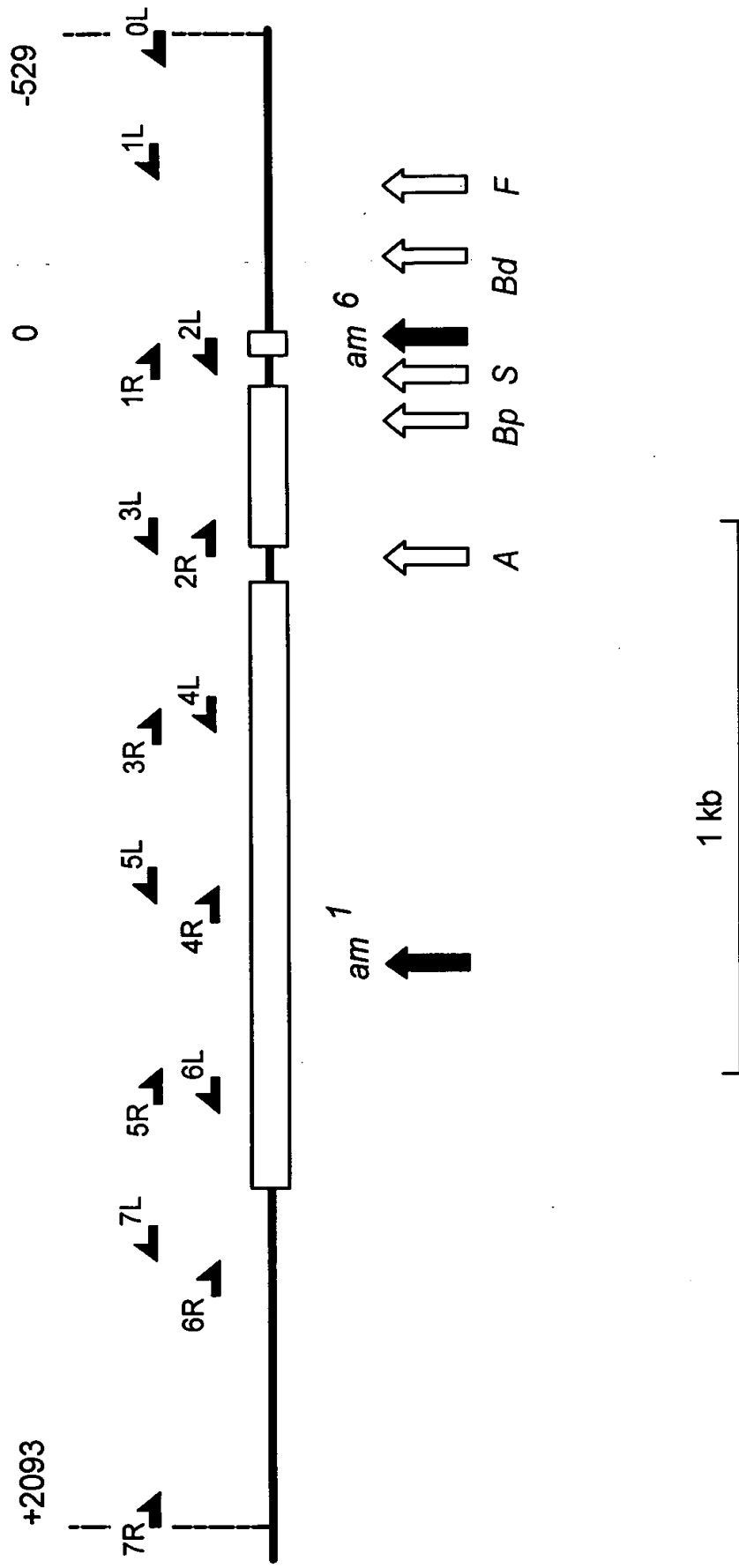


FIG. 12

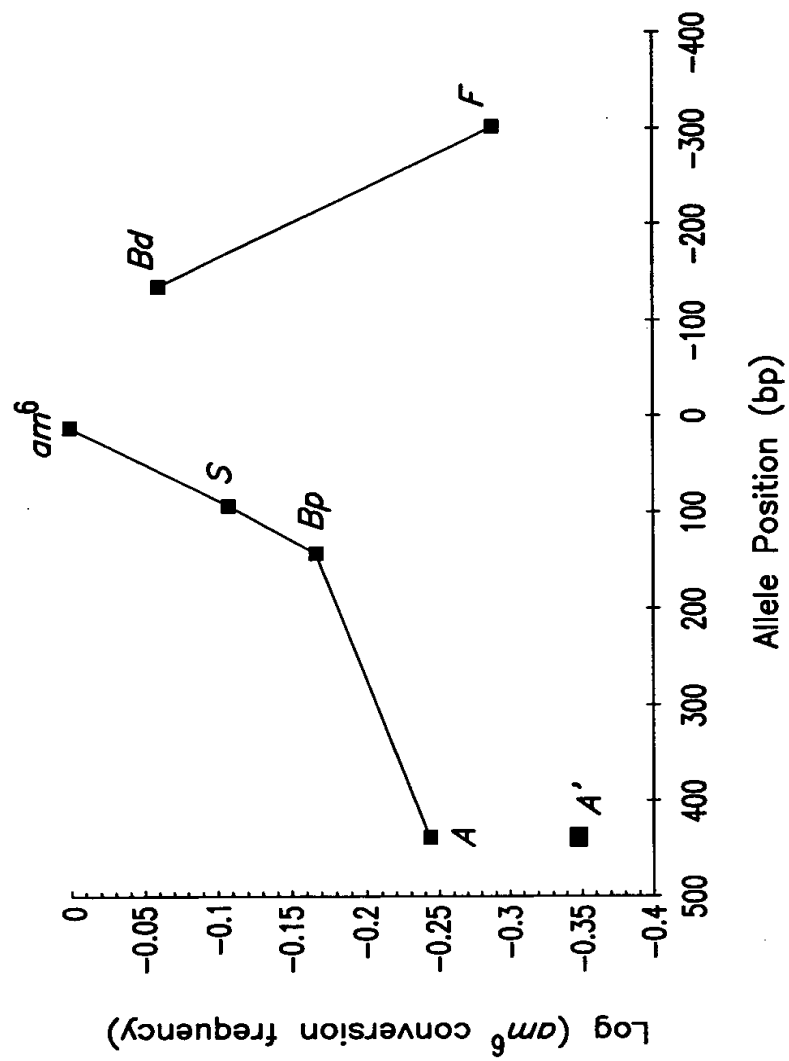
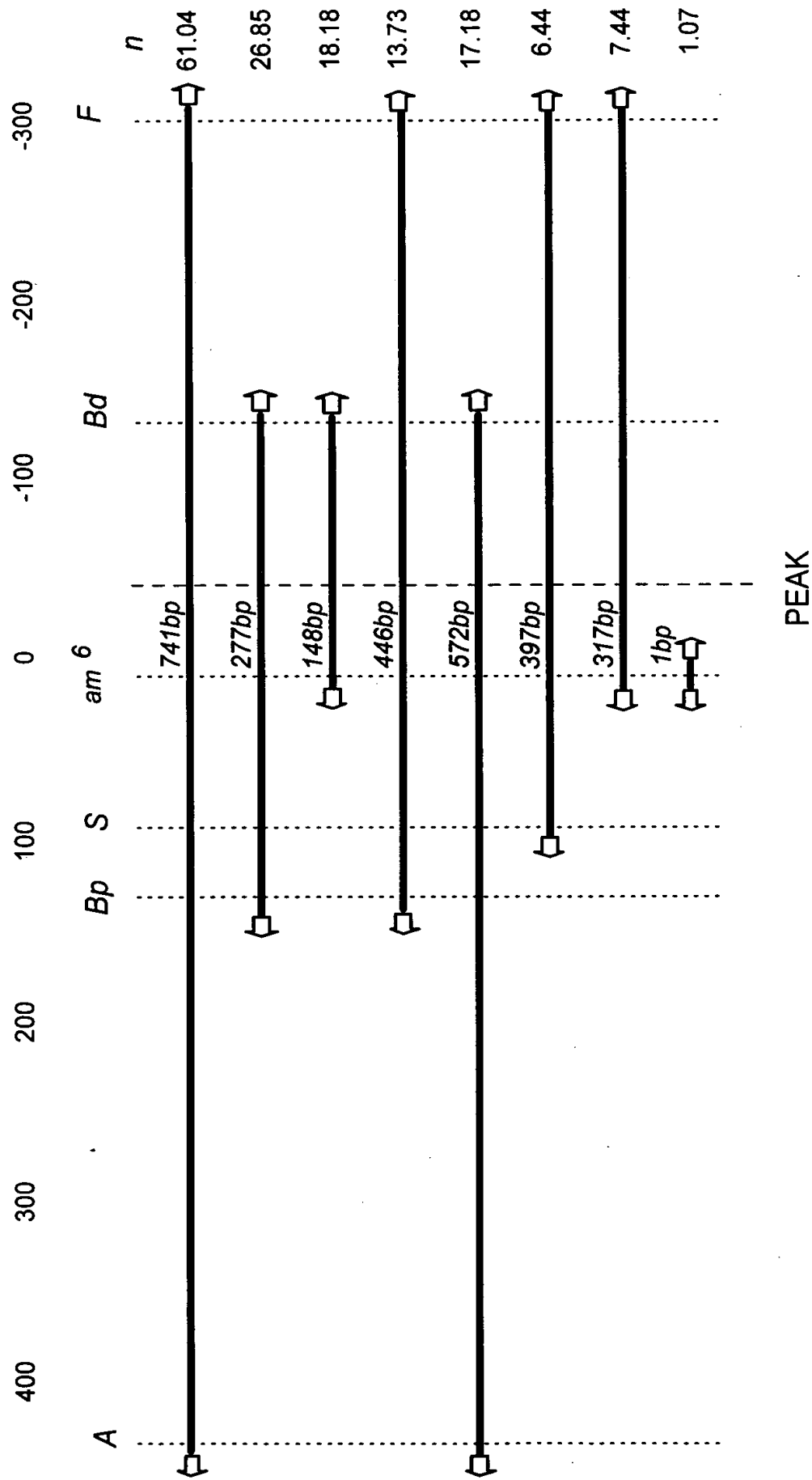


FIG. 13



# FIG. 14

